ENHANCING PHYSICAL-LAYER SECURITY IN WIRELESS POWERED COMMUNICATION NETWORK
CHALLENGES AND OPPORTUNITIES

Xing, Hong

Awarding institution:
King's College London

ENHANCING PHYSICAL-LAYER SECURITY IN WIRELESS POWERED COMMUNICATION NETWORK: CHALLENGES AND OPPORTUNITIES

HONG XING

KING'S COLLEGE LONDON

2015

ENHANCING PHYSICAL-LAYER SECURITY IN WIRELESS POWERED COMMUNICATION NETWORK: CHALLENGES AND OPPORTUNITIES

HONG XING

(*B. Eng., B. A. Zhejiang University*)

A THESIS SUBMITTED
FOR THE DEGREE DOCTOR OF PHILOSOPHY
IN
WIRELESS COMMUNICATIONS FACULTY OF NATURAL & MATHEMATICAL SCIENCES
KING'S COLLEGE LONDON

2015

Acknowledgements

I am sincerely grateful to my primal supervisor Prof. Arumugam Nallanathan for his knowledge of a wide range of physical-layer based resource allocations, unique perspective on novel research directions and above all strong support, without whom this thesis would have been impossible. I am especially thankful for all the opportunities that he has provided during my years of PhD studies, thanks to which I have benefitted immensely from having worked with collaborators from both within the UK and across the world. Among these experiences, my utmost gratitude goes to Dr. Rui Zhang, who shared with me his knowledge on the optimization for wireless communications, devotion to seeking blend of mathematical and physical principles in solving engineering problems, and meticulous attitude. Furthermore, I would like to express my appreciation to Prof. Kai-Kit Wong for his guidance throughout our collaboration with his vision, sagacity, and enthusiasm for innovation. With these role models' patience with me, my years of PhD studies have been such a rewarding and unforgettable experience.

In addition, I deeply appreciate the invaluable advice from Prof. Zhiguo Ding, one of our collaborators. I also feel very grateful to my former supervisor, Dr. Xiaoli Chu as well as the visiting scholar, Haijun Zhang, who both helped me at the very beginning stage of my PhD studies. In addition, I would like to acknowledge my second supervisor, Dr Mohammad Reza Nakhai, for his inspiring lectures and useful discussions with me.

Moreover, my heartfelt thanks go to Dr. Rui Zhang's group members in the Commun. & Network Lab of ECE department, National University of Singapore (NUS), who have made my visiting studies to NUS fruitful and enjoyable. Not only was I able to exchange ideas with these industrious colleagues in a stimulating work space, but I also developed genuine friendships with them inside and outside the lab, including but not exclusively Liang Liu, Suzhi Bi, Yong Zeng, Jie Xu, Yinghao Guo, Xun Zhou, Shixin Luo, Seunghyun Lee and Shuowen Zhang.

Furthermore, I have been fortunate to be constantly backed by my close friends who have talents in other related scientific subjects. In particular, the contributions to this thesis partially relied on intellectual and spiritual support from Dr. Yanwei Fu. I have also benefited a lot from consultation with Mr. Xiangyu Li who sheds light upon solving complex problems.

Lastly, I am thankful to my current and past colleagues here in the KCL's Centre for Telecommunications Research (CTR) lab: Yaqub Alwan, Yansha Deng, Adnan Aijaz, Changtao Zhong, Xinruo Zhang, Wan Wan Ariffin, Shuyu Ping, Zhenzhuang Miao, Christoforos Vlachos, Giorgos Chochlidakis, Seyed Ehsan Ghoreishi, and many others, who have convinced me that working hard is not incompatible with having fun.

Finally, I would like to dedicate this thesis to my beloved parents, Taisheng Xing and Xiaoying Zhao, for their selfless and boundless love all through the years. Their perpetual support is always there no matter how far I travel away from home, since the hearts of our family of three are intrinsically bound together.

Table of Contents

Abstract

Among various means of energy harvesting (EH) for green communications, radio-frequency (RF)-enabled wireless energy harvesting (WEH), inter alia, has recently drawn significant interest for its long operational distance and effective energy multicasting; it thus motivates the paradigm of wireless powered communication network (WPCN). Nevertheless, besides benefitting from the broadcast nature of wireless channels, WPCN is also vulnerable in terms of confidentiality and privacy of the data transmission, since legitimate information may be eavesdropped by unauthorized parties. To resolve this issue, physical-layer security (PLS) has been proposed as a promising solution to achieve information-theoretic security. This thesis is devoted to addressing some major challenges encountered in enhancing PLS for WPCN while exploiting opportunities gained from WPCN by pragmatic and prominent transmitting and/or cooperative strategies along with corresponding optimal (suboptimal) resource allocations.

This thesis begins with considering a three node single-input-single-output (SISO) fading wiretap channel, where the confidential messages sent to the information receivers (IRs) may be eavesdropped by the energy receivers (ERs) that are usually deployed nearer to the transmitter because of their high power receiving sensitivity. In this case, an artificial noise (AN)-aided transmission scheme, where the transmit power is split into two parts, to send the confidential message to the IR and an AN to interfere with the ER respectively, is proposed to facilitate the secrecy information transmission and yet meet the EH requirement. The fundamental challenges in balancing the goals between achieving PLS and satisfying ER's EH requirement are modeled by various secrecy performances versus harvested energy

trade-offs, the regions of which are enlarged by both dual decomposition-based optimal solutions and alternating optimization-based suboptimal solutions.

On the other hand, under circumstances where some ERs are trustful, their self-sustaining features can also be favourable to providing PLS by means of cooperative jamming (CJ). In the second part of the thesis, a novel harvest-and-jam (HJ) relaying protocol is proposed for multiple multi-antenna ERs to assist in the secrecy information transmission via one multi-antenna amplify-and-forward (AF) relay. Joint optimization of the CJ covariance and AF-relay beamforming is studied using semidefinite relaxation (SDR) under perfect and imperfect channel state information (CSI) respectively. In particular, for the imperfect CSI case, a novel approach that jointly models channel imperfections induced by an arbitrary number of CJ helpers is proposed to equivalently reformulate the worst-case robust optimization problem into the convex optimization framework.

Following the trend of WEH-enabled cooperative secrecy transmission, a more general wiretap channel with multiple WEH-enabled AF relays in the presence of multiple eavesdroppers all equipped with single antenna is studied in the last part of the thesis. To the end of combing the benefit of CJ and cooperative beamforming (CB), a new hybrid power splitting (PS) relaying strategy is proposed. In the first transmission phase each AF relay employs a PS receiver that splits a fraction of the received power for EH and consumes the rest for information receiving. In the second transmission phase the relay further divides its harvested power to forward the confidential information and to generate the jamming signals. The formulated secrecy rate maximization problems turn out to be very challenging due to the multiplicative variables in the relay weights. Under the centralized scheme, the global optimum joint CB and CJ solution is obtained for the static power splitting (SPS) case, while for the generalized dynamic power splitting (DPS) case, the global optimum CB-only solution is provided by utilizing SDR, which is then developed into a suboptimal joint CB and CJ design based on alternating optimization.

List of Tables

List of Figures

List of Abbreviations

AF	Amplify-and-Forward
AN	Artificial Noise
AP	Access Point
APC	Average Power Constraint
AWGN	Additive White Gaussian Noise
BCC	Broadcast Channel with Confidential Messages
CB	Cooperative Beamforming
CDI	Channel Distribution Information
CJ	Cooperative Jamming
CSCG	Circularly Symmetric Complex Gaussian
CSI	Channel State Information
CSIT	Channel State Information at the Transmitter
DF	Decode-and-Forward
d.o.f	Degree of Freedom
S-DPC	Secret Dirty-Paper Coding
EH	Energy Harvesting
ER	Energy Receiver
ESC	Ergodic Secrecy Capacity
EVD	Eigenvalue Decomposition
GSVD	Generalized Singular Value Decomposition
HJ	Harvest-and-Jam
IBCD	Inexact Block Coordinate Descent
ID	Information Decoding

IJ	Independent Jamming
IntRx	Integrated Receiver
IR	Information Receiver
KKT	Karush-Kuhn-Tucker
LMI	Linear Matrix Inequality
MF	Matched-Filter
MIMO	Multiple-Input Multiple-Output
MIMOME	Multiple-Input Multiple-Output Multiple-Eavesdropper
MISO	Multiple-Input Single-Output
MISOME	Multiple-Input Single-Output Multiple-Eavesdropper
MMSE	Minimum-Mean-Square-Error
MRC	Maximum Ratio Combining
MRT	Maximum Ratio Transmission
O-E	Outage-Energy
PDF	Probability Density Function
PLS	Physical-Layer Security
PPC	Peak Power Constraint
PS	Power Splitting
R-E	Rate-Energy
RF	Radio Frequency
RV	Random Variable
Rx	Receiver
SDP	Semidefinite Program
SDR	Semidefinite Relaxation
SIMO	Single-Input Multiple-Output
SINR	Signal-to-Interference-plus-Noise Ratio
SISO	Single-Input Single-Output
SNR	Signal-to-Noise Ratio
SOC	Second-Order Cone

SRM	Secrecy Rate Maximization
SVD	Singular Value Decomposition
SWIPT	Simultaneous Wireless Information and Power Transfer
TS	Time Switching
Tx	Transmitter
WIT	Wireless Information Transmission
w/o	Without
WPCN	Wireless Powered Communication Network
WPT	Wireless Power Transfer
w.r.t.	with respect to
w.l.o.g.	without loss of generality
ZF	Zero-Forcing

List of Notations

Throughout this thesis, scalars are denoted by lower-case letters, vectors denoted by bold-face lower-case letters, and matrices denoted by bold-face upper-case letters. Also, we define the following symbols:

$\boldsymbol{I}$	an Identity Matrix with Appropriate Dimension
$\mathbf{0}$	an All-Zero Matrix with Appropriate Dimension
$\boldsymbol{S}^{-1}$	the Inverse of the Square Full-Rank Matrix $\boldsymbol{S}$
$\mathrm{tr}(\boldsymbol{S})$	the Trace of the Square Matrix $\boldsymbol{S}$
$\boldsymbol{S} \succeq \mathbf{0}$	$\boldsymbol{S}$ is Positive Semi-Definite
$\boldsymbol{S} \preceq \mathbf{0}$	$\boldsymbol{S}$ is Negative Semi-Definite
$\boldsymbol{S} \succ \mathbf{0}$	$\boldsymbol{S}$ is Positive Definite
$\boldsymbol{S} \prec \mathbf{0}$	$\boldsymbol{S}$ is Negative Definite
$\boldsymbol{M}^T$	the Transpose of $\boldsymbol{M}$
$\boldsymbol{M}^\dagger$	the Conjugate of $\boldsymbol{M}$
$\boldsymbol{M}^H$	the Conjugate Transpose of $\boldsymbol{M}$
$\mathrm{rank}(\boldsymbol{M})$	the Rank of $\boldsymbol{M}$
$\mathbf{null}(\boldsymbol{M})$	the Null Space of $\boldsymbol{M}$
$\mathrm{vec}(\boldsymbol{M})$	a Column Vector Obtained by Stacking the Rows of $\boldsymbol{M}$ on Top of One Another
$\mathrm{vec}^{(-1)}(\boldsymbol{M})$	the Inverse Operation of vec(M)
$[\boldsymbol{M}]_{i,j}$	the (i,j)th Entry of $\boldsymbol{M}$
$diag(x_1, \ldots, x_N)$	a Diagonal Matrix with the Diagonal entries Given by $x_1, \ldots, x_N$
$[x_i]_{i=1}^N$	an $N \times 1$ Vector $\boldsymbol{x}$ with Each Element Indexed by x_i

$\mathcal{CN}(\boldsymbol{x}, \boldsymbol{\Sigma})$	the Distribution of a CSCG Random Vector with Mean Vector $\boldsymbol{x}$ and Covariance Matrix $\boldsymbol{\Sigma}$
$\sim$	"Distributed As"
$\xrightarrow{\text{a.s.}}$	"Almost Sure Converges to"
$\mathbb{C}^{x \times y}$	the Space of $x \times y$ Complex Matrices
$\mathbb{R}^{x \times y}$	the Space of $x \times y$ Real Matrices
$\|\boldsymbol{x}\|$	the Euclidean Norm of a Complex Vector $\boldsymbol{x}$
$\|\boldsymbol{x}\|.^2$	the Entry-Wise Absolute Value Square of a Complex Vector $\boldsymbol{x}$
$\mathbb{E}[X]$	the Statistical Expectation for Random Variable X
$\mathsf{Var}[X]$	the Statistical Variance for Random Variable X
$I(X;Y)$	the Mutual Information between Two Random Variables X and Y
$[x]^+$	$\max(0, x)$
$\boldsymbol{e}_k$	a Vector with its kth Component being 1, and All Other Components being 0
$\boldsymbol{A} \cdot \boldsymbol{B}$	the Product of Two Matrices
$\boldsymbol{A} \circ \boldsymbol{B}$	the Hadamard Product of Two Matrices
$\boldsymbol{A} \otimes \boldsymbol{B}$	the Kronecker Product of Two Matrices
$\|A\|$	the Cardinality of the Set A
$i.i.d.$	Independent and Identically Distributed
$P_r(\cdot)$	the Probability of an Input Random Event
$(\cdot)^*$	The Optimal Solution of the Input Optimization Variable

Chapter 1

Introduction

With the increasing demand for numerous amount of data traffic in wireless communication networks vis-a-vis the fact that mobile devices are usually constrained by their limited battery life and it is often costly to replace or recharge their batteries, energy harvesting (EH) has become attractive for realizing perpetual communications. Recently, radio-frequency (RF)-based wireless energy harvesting (WEH), which circumvents costly infrastructure and intermittence of the conventional means of EH, has therefore drawn an upsurge of interests owing to the development of RF-enabled circuits (see [3, 4] and the reference therein). With the transmit power, waveforms, and dimensions of resources etc., all being fully controllable, WEH-based technologies not only allow the wireless devices in low-power applications to scavenge energy from RF-signals that ubiquitously exist in wireless networks, but also enable "Shannon meets Tesla" [5], i.e., *simultaneous wireless information and power transfer (SWIPT)* over the same band of frequency, and therefore paves a new paradigm for truly self-sustaining next generation communications.

The advent of wireless networks, on the other hand, by permitting pervasive access to the on-line resources in its tremendous applications such as e-commerce, e-transaction, and cloud computing etc., has played a pivotal role in defining "the age of information". In comparison with WEH that essentially benefits from the broadcast nature of RF signals, wireless communications is nevertheless very susceptible to this intrinsic nature of wireless channels, since sensitive and confidential information, e.g., financial data, medical records, and customer files etc., is now easily exposed to unauthorised interception. Therefore, providing

privacy and integrity of data transmission has become indispensable in designing large-scale and heterogeneous wireless networks. In light of this, *physical-layer security (PLS)* has recently been proposed as a prominent solution to achieve wireless information-theoretic security [6], i.e., *perfect secrecy*, which is known to be a much stronger notion than computational security that is implemented by cryptography in upper layers. The key idea of PLS is to leverage the physical channel-induced randomness and impairments, such as channel fading and noise, to achieve both secure and reliable transmission for legitimate parties.

To accommodate the emerging *wireless powered communication network (WPCN)*, PLS needs to be carefully reexamined to take these two underlying "double sides of a coin" into account. This thesis is thus devoted to investigating some of the fundamental challenges in achieving PLS goals for WEH-enabled networks while fully exploiting opportunities brought by WPCN to enhance PLS. This chapter begins with the discussion on motivation of the thesis in typical application scenarios and then highlights the major contributions followed by the outline of the thesis.

1.1 Motivation

In WPCN, energy receivers (ERs) that are interested in harvesting energy from ambient RF signals are usually deployed in more proximity to the access point (AP) than information receivers (IRs) that intend to decode information, to meet their different receiving power sensitivity (-10dBm for ERs versus -60dBm for IRs). As such, albeit being able to facilitate SWIPT, WPCN gives rise to a potential threat to information-theoretic security: what if the ERs attempt to eavesdrop the information destined to IRs rather than harvest energy as presumed to? On the other hand, PLS issues in the emerging cooperative communications, such as heterogeneous networks, device-to-device systems, and relaying networks etc., have been widely studied over the recent years by exploiting cooperative secure transmission schemes, particularly, *cooperative jamming (CJ)* [7, 8], which exploits friendly jamming to degrade eavesdropper's decoding ability. However, the benefits of CJ would be

quite compromised if the potential helpers are unwilling to cooperate due to their own limited battery capacity. In this regard, WPCN that makes self-sustainable communication possible also provides opportunities to circumvent this ultimate bottleneck in energy-constrained cooperative networks.

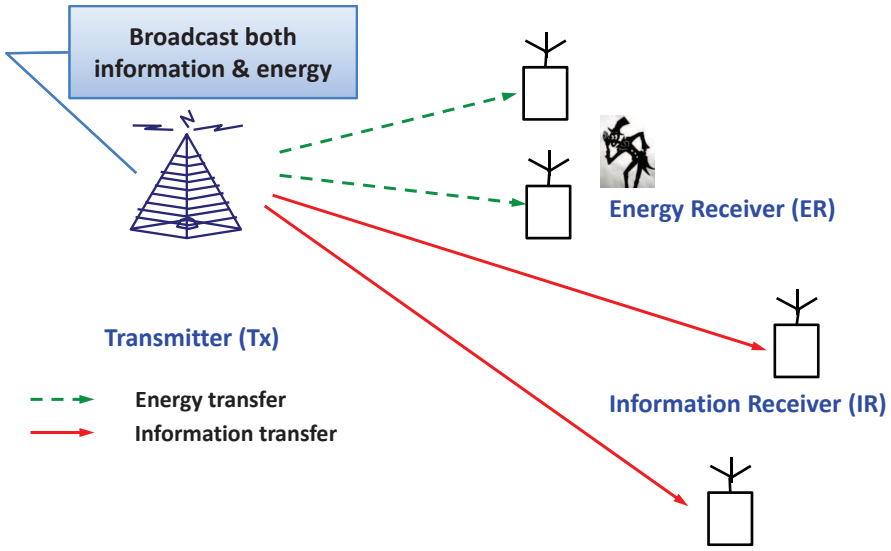


Figure 1.1: A SWIPT system with separate IRs and ERs deployed "far" and "near" to the AP, respectively.

A typical SWIPT system with separate IRs and ERs is shown in Fig. 1.1, where a hybrid AP with constant power supply broadcasts RF signals to a set of distributed user terminals scheduled in a "near-far" fashion aforementioned. In view of secrecy information transmission to the IRs, it is easy to identify two conflicting goals in the transmission design: the power of the information signal at the energy receiver (ER) is desired to be made large for efficient EH, but also needs to be kept sufficiently small to prevent information leakage. To resolve this conflict, transmit power has to be split in part for artificial noise (AN) to interfere with the ER. It is worthy of noting that, although soliciting for AN in secret communications has been extensively studied in the literature, the AN design in SWIPT is fundamentally different since it also contributes to the total power harvested at the ER, which is otherwise beneficial for satisfying the EH requirement.

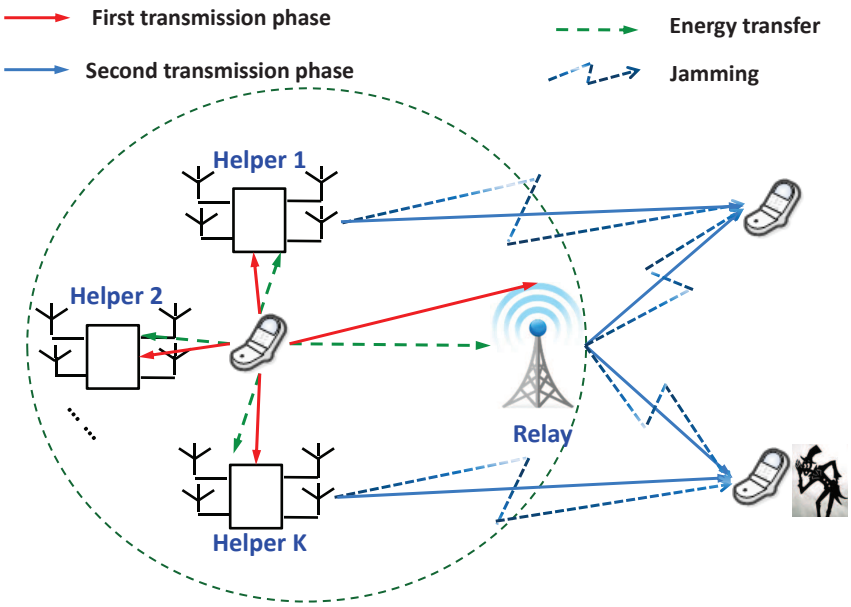


Figure 1.2: A cooperative SWIPT system with CJ-aided multi-AF relaying.

Unlike the above situation where the ERs in the SWIPT systems attempt to intercept the confidential information for the IR, it is possible that some of the WEH-enabled ERs are actually cooperative. For example, following the recent advances in WPCN, in addition to conventional cooperative schemes, such as amplify-and-forward (AF) and decode-and-forward (DF), helpful ERs relieved of energy-constrained concerns in a cooperative SWIPT network can now work as self-sustaining friendly jammers who protect the forwarded confidential information against eavesdropping in the second transmission phase using its harvested energy from the first transmission phase as depicted in Fig. 1.2.

First, considering the case where the friendly ERs and the information-forwarding relay are separately located, an immediate challenge vis-a-vis the scenario shown in Fig. 1.1 is that the eavesdroppers now incline to conceal from the transmitter (Tx), since they do not need to assist the Tx in estimating their channels for efficient WEH. Consequently, robust transmission design is motivated so as to combat the imperfect channel state information (CSI)

regarding the eavesdroppers. Furthermore, it is worth noting that most of the well studied CJ-aided cooperative schemes in the literature cannot be directly applied to this case of interest, since the power allocations in WPCN is more challenging than that with constant power supply in the sense that presently the design of AN is also subject to their respective EH channels. Second, considering a more general case where the friendly ERs and the information-forwarding relays are co-located, i.e., the available power at each relay is split into two parts: one for forwarding the confidential message and the other for CJ, the secrecy performance with the aid of WEH-enabled relays is expected to considerably outperform the CJ and *cooperative beamforming (CB)* separately designed ones, thanks to the vast degree-of-freedom (d.o.f) brought about by relays. However, due to the practical circuit limit that the RF signals cannot be decoded during its energy being harvested, *dynamic power splitting (DPS)* that is known to be the best possible receiver architecture for SWIPT so far [9], has to be employed at the relays, which poses intractable non-convexity to the joint optimization of CB and CJ due to the multiplicative relay weights. As a result, current state-of-art secrecy rate maximization (SRM) algorithms that jointly optimize the power splitting (PS) ratios and relay beamforming in the literature only converge to local optimum solutions.

1.2 Contributions of the Thesis

The major contributions of this thesis are summarized as follows.

Considering a simplified three-node single-input single-output (SISO) fading wiretap channel, an AN aided transmission scheme is proposed in Chapter 3 to facilitate the secrecy information transmission to IRs and yet meet the EH requirement for ERs. Problems are formulated to minimize the outage probability for the IR for delay-limited secrecy transmission, or to maximize the ergodic secrecy capacity (ESC) for the IR for no-delay-limited secrecy transmission, subject to combined average and peak power constraints at the Tx as well as an average EH constraint at the ER. The formulated problems, however, are shown to be both

non-convex. For each of the two problems, a dual decomposition based method is first proposed to solve it optimally and then an efficient suboptimal algorithm is designed by iteratively optimizing the transmit power allocations and power splitting ratios over different fading states. Finally, the proposed schemes are evaluated by various trade-offs for wireless (secrecy) information transfer versus wireless power transfer (WPT).

In a multi-antenna AF relay wiretap channel in the presence of Alice, Bob and Eve all equipped with single-antenna, an innovative multi-antenna *harvest-and-jam (HJ)* relaying protocol is proposed in Chapter 4. The contributions of this chapter are threefold. First, with perfect CSI, in addition to the joint optimal solutions, two near-optimal schemes with much reduced complexity are proposed. One of the schemes exploits the optimal structure of the relay weight matrix, which is a novel extension of a similar relay beamforming matrix optimal to the two-way relay channel [10], while for the other *null-space jamming*, a semi-closed form solution for the relay weight matrix is provided. Second, besides the imperfect eavesdroppers channel, legitimate channels such as those from the HJ helpers to the legitimate receiver (Rx) and from the AF relay to the Rx are jointly modeled with imperfect CSI, and multiple semi-indefinite non-convex constraints induced have been, for the first time to the best knowledge of the author, equivalently replaced by linear matrix inequalities (LMIs) in order to fit in with the convex optimization. Third, a rank-one reconstruction algorithm to enable transmit beamforming has been proposed to provide promising performance by exploiting the structure of the semi-definite relaxation (SDR)-based solutions.

A CJ-aided wiretap channel with multiple WEH-enabled AF-operated relays in the presence of multiple single-antenna eavesdroppers is studied in Chapter 5. The achievable secrecy rates are maximized subject to individual EH power constraints of relays by jointly optimizing the CB and CJ covariance matrices using the technique of SDR. In the centralized case with global CSI, optimal AF relay beamforming is provided in closed-forms, respectively, for *static power splitting (SPS)*-enabled AF

relays performing joint optimization of CB and CJ, and DPS-enabled AF relays performing only CB. The optimal solutions are derived based on rigourous proof for the tightness of SDR. In particular, the global optimum AF relay beam along with its DPS ratios without (w/o) the use of CJ has been derived for the first time as far as the authors know, which thus gives a tight upper-bound for wireless-powered CB design in a wide range of SWIPT-enabled applications. For joint optimization of CB and CJ with DPS-enabled AF relays, a viable suboptimal algorithm that iteratively obtains the optimal AN beams, relay beams as well the power splitting ratios is proposed based on alternating optimization.

1.3 Outline of the Thesis

The thesis is organized as follows. Some basic concepts of information-theoretic security are introduced in Chapter 2, followed by an overview of the state-of-art signal processing techniques achieving wireless PLS in a variety of wiretap channels, and the recent advances in enhancing PLS for WPCN are summarized. The main contributions are presented in Chapters 3-5. Specifically, in Chapter 3, the potential security issues induced by SWIPT in a three-node SISO fading wiretap channel are discussed and the compromise of information-theoretic security on the system EH requirements is characterized by various trade-offs, such as (secrecy) outage-energy (O-E) trade-off and (secrecy) rate-energy (R-E) trade-off. Chapters 4 and 5, on another front, focus on opportunities gained by WPCN for PLS designs. In Chapter 4, (worst-case robust) SRM problems are considered for a cooperative SWIPT network in the presence of one eavesdropper by exploiting wireless powered CJ, and the AN covariances and AF beamforming are jointly optimized under the perfect and imperfect CSI, respectively, while in Chapter 5, the above model is generalized to a more sophisticated case with multi-AF relays in the presence of multiple eavesdroppers by investigating both wireless powered CJ and CB, and dedicated hybrid PS receiver at each relay is employed. Finally the thesis is concluded and some interesting directions for future studies are pointed out in Chapter 6.

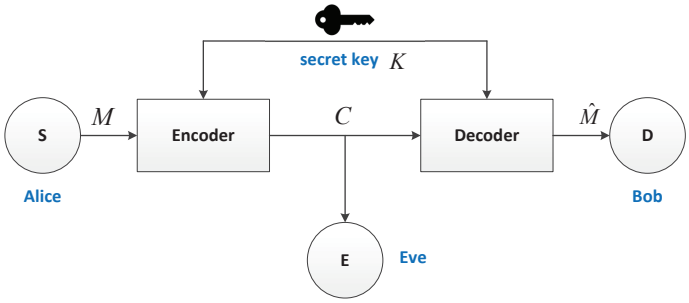
Chapter 2

Literature Review

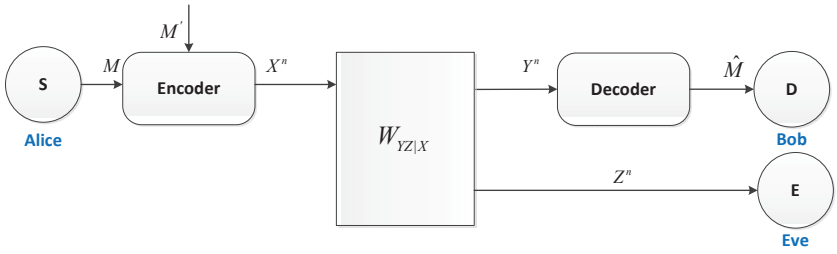
In this chapter, a unified literature review on the related work is presented. In Section 2.1, fundamentals of information-theoretic secrecy are briefly introduced, which provide the fundamental limits of various PLS enhancements schemes. It is followed by Section 2.2, where an overview of the state-of-art signal processing techniques applied in a variety of wiretap channels is given. Finally, the latest progress in secrecy communications with the upsurge of SWIPT technologies is summarized in Section 2.3.

2.1 Fundamentals of Information-theoretic Secrecy

Information-theoretic secrecy not only allows for precise and quantitative analysis of various secrecy schemes across different layers communication system, but also provides fundamental limits of the transmission rate over which the transmitter is able to achieve reliable and secure communication with the legitimate Rx. The objective of this section is to introduce several information-theoretic secrecy metrics along with their operational meaning and to present the classical model first proposed by Wyner [11], which is, albeit basic, amendable to more complex wiretap channels investigated in the literature (see Section 2.2 for more detail). Furthermore, the coding mechanism that achieves information-theoretic secrecy based on this model is intuitively interpreted.



(a) Shannon's cipher system.



(b) Wyner's wiretap channel.

Figure 2.1: Classical information-theoretic secrecy model.

2.1.1 Information-Theoretic Secrecy Metrics

To illustrate the significant notion of information-theoretic secrecy, *perfect secrecy* [12], Claude Shannon's classical model of a cipher system as seen in Fig. 2.1(a) needs to be recalled. The objective of this cipher system is to reliably transmit a confidential message M from the transmitter (Alice) to a legitimate receiver (Bob) in the presence of an eavesdropper (Eve), who intercepts all transmitted signals. In this model, the cryptographic encoder-decoder that maps the plain text M to the cipher text C and vice versa is realized by sharing a random key K that is exclusively between Alice and Bob. *perfect secrecy* is achieved if the confidential message M is statistically independent of C, i.e., $p(m) = p(m|c)$, where m and c are instances of M and C, respectively. According to basic inequalities from information theory, it is equivalent to state that the mutual information between M and C, $I(M;C)$, is identical to zero. In other words, assuming M is uniformly distributed over a message set $\{1,\ldots,2^N\}$, the only way left behind for Eve is to guess M with the successful probability of 2^{-N}.

Remark 2.1.1. *It is worthy noting that perfect secrecy is a much stronger notion than those used in computational security, since it is a quantitative measure independent of any assumption regarding Eve's computational capacity and therefore ensures that C intercepted by Eve reveals no information about M. It was shown that it is achievable using a simple one-time pad encryption [13], which induces a secret key of the same length as the message itself.*

Perfect secrecy defined by Shannon that requires exact independence between M and C is too stringent to be implemented in practice, since the advantage of perfect secrecy is considerably offset by generation of long-bit keys as above mentioned. Consequently, as is often done in information-theoretic secrecy, perfect secrecy in Shannon's cipher system is relaxed by only asking for statistical independence between M and K in an asymptotic sense, which necessitates a metric to measure the statistical dependence between M and K. Among many ways to define such metrics, the most commonly used one is *strong secrecy* [14] given by

$$\lim_{n\to\infty} I(M;C^n) = 0. \tag{2.1}$$

In comparison, *weak secrecy* measures the rate at which the information about M is leaked in C^n, which is given by

$$\lim_{n\to\infty} \frac{1}{n} I(M;C^n) = 0, \tag{2.2}$$

where C^n denotes a length-n codeword encoding M.

2.1.2 Wyner's Wiretap Channel

It is important to note that Shannon's cipher system cannot be applied directly to the PLS system of the interest, especially to the wireless PLS system, since the intrinsic property of (wireless) communication system that is not captured in Shannon's model is the presence of impairments in the communication channel, and therefore his model is restricted to the assumption of perfect interception of

the transmit symbols. However, the significant motivation for PLS is to explicitly consider the imperfection induced by communication channels. The basic model for PLS, called the *wiretap channel*, was pioneered by Wyner [11], which captured the joint problem of reliable and secret communication over noisy channels. As illustrated in Fig. 2.1(b), the objective is that Alice is able to reliably and secretly communicate with Bob at rate R, by encoding messages $M \in \{1; 2^{nR}\}$ into codewords X^n of length n and transmitting X^n over a noisy memoryless broadcast channel, characterized by a transition probability $W_{YZ|X}$. Y^n and Z^n denote the receiver observation at Bob and Eve, respectively, and $M' \in \{1; 2^{nR'}\}$ is a local random number generator (each of its realization is known only by Alice) to assist in encoding of message M. In this model, R is *achievable* in the sense that there exists a codebook, called *wiretap code* with increasing block length n such that

$$\lim_{n\to\infty} P_r(\hat{M} \neq M) = 0, \ \lim_{n\to\infty} I(M; Z^n) = 0, \tag{2.3}$$

which guarantees reliability and secrecy, respectively.

Remark 2.1.2. *It is worthy of noting that the wiretap channel model introduced by Wyner is different from Shannon's cipher system, besides introducing noise, also in the respect that it does not include any shared secret key between Alice and Bob, and therefore allows for independent design from cryptography. Furthermore, one fundamental difference between the wiretap code and the conventional communication code lies in the* stochastic encoder *M', the role played by which will be clarified shortly. Intuitively, one can think of the function of M' analogous to the key used in a one-time pad, because both of them aims for randomizing the source message M. In addition, the wiretap channel model herein assumes that Alice and Bob know the Eve's channel perfectly, which might be too strong to account for a passive eavesdropper. However, recent advances in signal processing have resolved this issue in part that will be detailed in Section 2.2.*

The supremum of all achievable R defines one of the information-theoretic

metrics for PLS, i.e., *secrecy capacity*, which is shown to be [11, 15]

$$C_s = \max_{V \to X \to YZ} (I(V;Y) - I(V;Z)), \tag{2.4}$$

where V is an auxiliary random variable (RV).

The structure and design of wiretap code that achieves C_s is quite nontrivial. To give an intuitive illustration of the coding mechanism behind (2.4), a simple model with noiseless main channel but noisy eavesdropper's channel is considered. Since reliable communication is automatically achieved under this assumption, it is safe to consider $V = X$ without loss of generality (w.l.o.g.). Now assuming that given the joint distribution of this broadcast channel, $W_{YZ|X}$, the achievable rate of the respective channel between Alice to Bob and Eve, are $R + R'$ and R', it follows from (2.4) that a positive secrecy rate $R > 0$ is achievable since Bob's channel is less noisy. Furthermore, for a noiseless channel with binary input, it is easy to verify that $\frac{1}{n} I(M; Y^n) \leq 1$, which ensures that $R + R' \leq 1$,

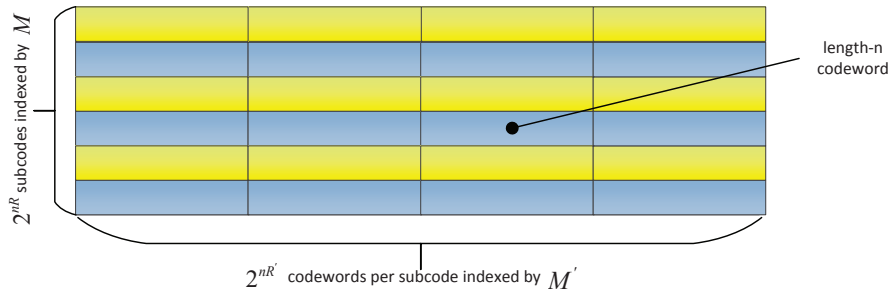


Figure 2.2: Nested structure of a wiretap code [1, Figure 1.3].

In the following, a *nested-structure* wiretap code using stochastic encoding is constructed as depicted in Fig. 2.2. Specifically, given a message set represented by $\{1; 2^{nR}\}$ with R defined earlier, each of the 2^{nR} distinct messages is randomly mapped into a codeword indexed by M' from a $(2^{nR'}, n(1-R))$ codebook, which is designed to be *capacity-achieving* for the eavesdropper's channel. The codebook corresponding to one particular message forms a bin, which is a *subcode* of the wiretap

code, and consequently there are totally 2^{nR} subcodes indexed by M. Now it is ready to evaluate the information leaked about the message M. First, $I(M;Z^n)$ is rewritten by manipulating with basic information-theoretic identities as follows.

$$I(M;Z^n) = I(X^n;Z^n) - H(M') + H(M'|MZ^n), \tag{2.5}$$

where $I(X^n;Z^n)$ represents the information leaked about the codewords, $H(M')$ represents the entropy of the local random generator, and $H(M'|MZ^n)$ measures Eve's uncertainty within the subcode. It is intuitively to observe from (2.5) that if $H(M')$ is large enough to compensates the information leaked about the codeword X^n and the uncertainty induced by $H(M'|MZ^n)$, the information leakage about the message tends to vanish. Fortunately, since the subcode $(2^{nR'}, n(1-R))$ is designed to be capacity-achieving, which is possible since $R'+R \leq 1$, it can be shown by [16] that $\frac{1}{n}H(M') \approx \frac{1}{n}I(X;Z^n)$ and $\frac{1}{n}H(M'|MZ^n) \approx 0$ are guaranteed so that $\frac{1}{n}I(M;Z^n) \approx 0$. Combining with $P_r(\hat{M} \neq M) = 0$ under the noiseless main channel, this approach of code construction is shown to achieve secrecy rate R with (2.3) satisfied in a weak sense.

Note that more detailed discussion on capacity-achieving wiretap code design is out of the scope of this thesis. However, the random coding mechanism introduced above also sheds light upon the design of wiretap code in other wiretap channels, the principle of which will be favourable in appreciating the fundamental results reviewed in the sequel. On the other hand, there is another information-theoretic secrecy branch for PLS dealing with *secret-key generation* that focuses on the distillation of secrecy from common randomness by public discussion over noiseless side channel of unlimited capacity. The related discussion is out of the scope of this paper and it is mentioned herein nevertheless to emphasize that PLS can complement, rather than replace, existing cryptographic techniques used in upper layers, for instance, by providing a secure means of randomness sharing for generating secret keys.

2.2 An Overview of Signal Processing Approaches for Improving PLS

Compared with the immediate benefit brought by multiple antennas and/or multiple users to reliable communications w/o secrecy concerns, the advantage of extra d.o.f could have been severely compromised by the simultaneously increased decoding capacity of potential eavesdroppers w/o careful design. In this section, an overview of some typical approaches for achieving gains in secrecy diversity and/or multiplexing by judiciously designing the extra dimensions of resources is given.

The section commences with a presentation of transceiver design algorithms for the classical three-user multiple-input multiple-output (MIMO) wiretap channel. The work generalized to multi-Eve wiretap channel is expected to be more arduous, because increasing the number of eavesdroppers indicates the increasing interception ability of unauthorized parties even if they do not collude. A critical countermeasure that paves a new way for effective transceiver design is to generate *artificial noise (AN)*, i.e., synthetic noise embedded in conjunction with the confidential messages, which also refers to *cooperative jamming (CJ)* when it is generated by separate helpers in cooperative communications. Subsequently, the fading wiretap channel is examined, the intrinsic randomness of which that used to be regarded as a downside for wireless communications, however, plays a major role in achieving secret communications. Lastly, another interesting wiretap channel worth investigating is the relay wiretap channel, where a joint optimization of precoders in both relays and the Tx is an effective enabler in enhancing secrecy.

In each subsection, key techniques of secrecy transceiver designs in the literature are summarized under both perfect and imperfect (or even no) channel state information at the transmitter (CSIT), the latter of which necessitates robust transmission strategies.

2.2.1 MIMO Wiretap Channels

A classical type of wiretap channel, three-user MIMO wiretap channel as shown in Fig. 2.3 consists of a transmitter (Alice), a legitimate receiver (Bob), and an eavesdropper (Eve), each of which is equipped with N_T, N_R and N_E number of antennas, respectively. The received signal by the legitimate receiver is given by

$$\boldsymbol{y}_b = \boldsymbol{H}_b \boldsymbol{x}_a + \boldsymbol{n}_b, \tag{2.6}$$

whilst that also received at the eavesdropper is

$$\boldsymbol{y}_e = \boldsymbol{H}_e \boldsymbol{x}_a + \boldsymbol{n}_e, \tag{2.7}$$

where $\boldsymbol{x}_a \in \mathbb{C}^{N_T \times 1}$ is the transmit signal with covariance matrix $\mathsf{E}[\boldsymbol{x}_a \boldsymbol{x}_a^H]$ denoted by $\boldsymbol{Q}_x$; $\boldsymbol{H}_b$ and $\boldsymbol{H}_e$ represent the complex MIMO channels from Alice to Bob and Eve, respectively; $\boldsymbol{n}_b$ and $\boldsymbol{n}_e$ are additive white Gaussian noise (AWGN) at the receivers of Bob and Eve, respectively.

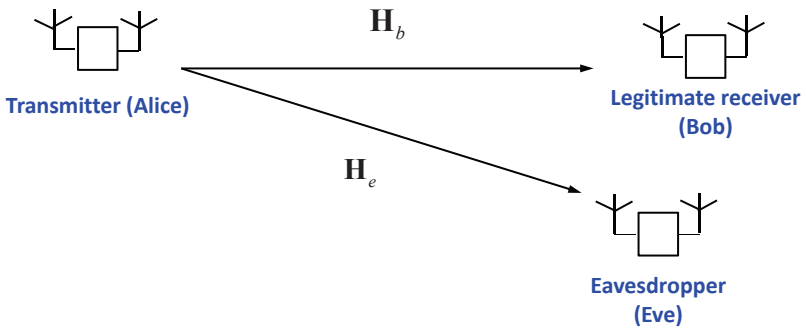


Figure 2.3: A three-user MIMO wiretap channel.

Note that the wiretap channel model described above is also referred to as *multiple-input, multiple-output, multiple-eavesdropper (MIMOME)* channel [17], a special class of which, *multiple-input, single-output, multiple-eavesdropper (MISOME)* channel [18], corresponds to the case where Bob is equipped with

only one single antenna. As MIMOME suggests, the "multi-eavesdropper" herein comprises the scenario when there are multiple antennas equipped on a single eavesdropper or when there is a group of geographically dispersed but colluding single-antenna eavesdroppers. Note that the case in which there are several non-colluding eavesdroppers is referred as *compound* wiretap channel [19, 20].

Assuming AWGN at the receivers, the secrecy capacity under an average power constraint (P) for the transmit covariance ($\boldsymbol{Q}_x$), i.e., $\mathsf{Tr}(\boldsymbol{Q}_x) \leq P$, is given by [18, 21–23]

$$C_{sec} = \max_{\boldsymbol{Q}_x,\, \mathsf{Tr}(\boldsymbol{Q}_x)\leq P} [I(\boldsymbol{X}_a; \boldsymbol{Y}_b) - I(\boldsymbol{X}_a; \boldsymbol{Y}_e)]. \tag{2.8}$$

For Gaussian input signaling, which is the optimal choice for achieving the secrecy capacity given in (2.8), (2.8) is simplified as

$$C_{sec} = \max_{\boldsymbol{Q}_x,\, \mathsf{Tr}(\boldsymbol{Q}_x)\leq P} [\log\det(\boldsymbol{I} + \boldsymbol{H}_b \boldsymbol{Q}_x \boldsymbol{H}_b^H) - \log\det(\boldsymbol{I} + \boldsymbol{H}_e \boldsymbol{Q}_x \boldsymbol{H}_e^H)], \tag{2.9}$$

where it is assumed that $\boldsymbol{n}_b$ ($\boldsymbol{n}_e$) is an AWGN denoted by $\boldsymbol{n}_b \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_{N_R})$ ($\boldsymbol{n}_e \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_{N_E})$) for simplifying analysis.

1) Perfect CSI

Not very surprisingly, even for the case when instantaneous CSI for both Bob and Eve are perfectly known at the transmitter, a precise characterization of the three-user MIMO wiretap channel under an average power constraint of $\mathsf{Tr}(\boldsymbol{Q}_x) \leq P$, in general, remains an open problem. However, in some special cases, closed-form solutions can be found. For example, in the MISOME case where $N_R = 1$, N_T, $N_E > 1$, the optimal transmit beamforming is given by $\boldsymbol{Q}_x = P\boldsymbol{\psi}_m \boldsymbol{\psi}_m^H$, where $\boldsymbol{\psi}_m$ is the normalized generalized eigenvector corresponding to the largest generalized eigenvalue λ_m as follows.

$$(\boldsymbol{I} + P\boldsymbol{h}_b \boldsymbol{h}_b^H)\boldsymbol{\psi}_m = \lambda_m (\boldsymbol{I} + P\boldsymbol{H}_e^H \boldsymbol{H}_e)\boldsymbol{\psi}_m. \tag{2.10}$$

Another special circumstance is that when $\boldsymbol{Q}_x$ is known to be full rank [24, *Theorem 2*]. For the general unsolvable MIMOME case, nevertheless, the asymptotically optimal solution in the high signal-to-noise ratio (SNR) regime can be found by decomposing the system into parallel channels based upon the generalized singular value decomposition (GSVD) of the matrix pair $(\boldsymbol{H}_b, \boldsymbol{H}_e)$ [17]. The optimal power allocation for GSVD transmission precoding was derived in [25] and shown to empirically achieve the MIMO secrecy capacity given in (2.9).

An alternative type of power constraint, namely, a matrix power covariance constraint, $\boldsymbol{Q}_x \preceq \boldsymbol{S}$, was considered in [26], where the MIMO wiretap channel was reexamined by exploiting the derivative relationship between mutual information and mean-squared error to provide a closed-form expression for the secrecy capacity, which was shown to be

$$C_{sec}(\boldsymbol{S}) = \sum_{i=1}^{\lambda} \log \alpha_i. \tag{2.11}$$

(2.11) is achieved when $\boldsymbol{Q}_x = \boldsymbol{S}$, where α_i's are the generalized eigenvalues of the pencil

$$(\boldsymbol{S}^{\frac{1}{2}} \boldsymbol{H}_b^H \boldsymbol{H}_b \boldsymbol{S}^{\frac{1}{2}} + \boldsymbol{I}, \boldsymbol{S}^{\frac{1}{2}} \boldsymbol{H}_e^H \boldsymbol{H}_e \boldsymbol{S}^{\frac{1}{2}} + \boldsymbol{I}) \tag{2.12}$$

that are greater than one.

The secrecy capacity given in (2.11), for which the imposed matrix power constraint places considerable limits on the per-antenna power power and the transmit correlation structure, is expected to achieve inferior performance to that under an average power constraint given in (2.9). The relationship between (2.11) and (2.9) is characterized by [27, *Lemma 1*]

$$C_{sec}(P) = \max_{\boldsymbol{S} \succeq \boldsymbol{0}, \mathsf{Tr}(\boldsymbol{S}) \leq P} C_{sec}(\boldsymbol{S}), \tag{2.13}$$

where for any given $\boldsymbol{S}$, $C_{sec}(\boldsymbol{S})$ is computed as in (2.11).

The fundamental work on information-theoretic security pioneered by Wyner [11], Csiszár and Köner [15], have already shown that a positive secrecy capacity can be achieved only if the eavesdropper's channel is a degraded version of the main channel. Recently, a variety of physical-layer techniques from signal processing perspective were devoted to breaking through this assumption, among which, exploiting synthetic interference (from the co-located transmitter, known as AN; from an external helping interferer, referred as CJ) to deteriorate Eve's interception capability has been shown effective. Although multi-antenna precoding embedded with AN is regarded promising, for a three user MISOME wiretap channel, it has been shown in [18, 22] that if Eve's channel is known at the transmitter, the secrecy capacity is achievable w/o transmission of AN.

2) Imperfect CSI

On the other hand, considering an extreme case where no eavesdroppers CSI is available at all, well-known isotropic AN [21] is designed to effectively degrade the Eves channel, nevertheless nulled out at the intended receiver. Therefore, it is significant to investigate the role of AN with exposure to different level of Eves CSI. The assumption with regards to eavesdroppers' imperfect CSI can be mainly categorized into three classes: the statistical distribution of $\boldsymbol{H}_e$, a known bound for the error where the uncertainty in knowledge of $\boldsymbol{H}_e$ lies within, and no Eve's CSI at all. For the first case, the AN injection strategy proposed by Goel and Negi [21] remains the best known secure transmission strategy that optimizes the number of spatial dimensions and power allocated to the AN. For the second case with norm-bounded error, [28] studied the optimal transmit strategy by exploiting the relationship between the multi-input single-output (MISO) wiretap channel in the presence of one single-antenna eavesdropper and the cognitive radio MISO channel. Huang et al. in [29] considered the robust transmit designs against Eve's bounded channel mismatches with the aid of AN generated by an external helper, also termed as *cooperative jamming (CJ)* for similar MISO wiretap channel. Robust transmit covariance matrices were obtained therein with and w/o CJ, respectively, by

transforming the nonconvex max-min problem into a quasiconvex one based on the worst-case secrecy rate formulation. It was further concluded in [29] that the robust transmit designs with CJ is particularly helpful when the Eve's channel is imperfect, although the benefit of CJ is not seen under perfect Eve's CSI. For the third case, since CSI of Eve(s) are totally unknown, nothing might be more intelligent than transmitting as much as possible spatially isotropic AN to the nullspace of the desired signal [30, 31]. Specifically, [30] focused on maximizing the amount of power allocated to the AN that hides the confidential information from a potential eavesdropper while guaranteeing a prescribed signal-to-interference-plus-noise ratio (SINR) at the legitimate Rx assuming that Eve employs the optimum beamformer at its receiver. In addition, robust coding schemes from the perspective of information-theoretic security can also be designed to circumvent eavesdropping irrespective of the Eve's CSI, which is beyond the discussion of this thesis.

Besides, instead of fixing AN in the nullspace of the legitimated channel[29, 32], [33] considered the optimal AN-aided transmit strategy that simultaneously optimizes the transmit and AN covariances to maximize the secrecy rate. These joint optimization approaches under both perfect and imperfect CSI case obfuscate eavesdroppers's interception more effectively in the sense that the AN can now take on any spatial pattern rather than a specific kind, for example, spatially isotropic AN [34], beamforming AN [35, 36], which are all suboptimal in terms of secrecy rate maximization.

2.2.2 Fading Wiretap Channels

Multi-path fading phenomenon in wireless communications, albeit negative in most wireless communications, is nevertheless beneficial to achieving nonzero ergodic secrecy rate or outage secrecy rate [6, 37]. Before examining fundamental results for fading wiretap channel in literature, a three-node fading wiretap channel that is corrupted by multiplicative fading in addition to AWGN is introduced to illustrate the definition of ESC and secrecy outage probability. As shown in Fig. 2.4, the source

S communicates with the destination D in the presence of an eavesdropper, E, over the main channel, denoted by g_M, and the eavesdropper channel, denoted by g_E, respectively. The power gains experiencing fading of the two complex channels are correspondingly denoted by $h_M = |g_M|^2$ and $h_E = |g_E|^2$. The *i.i.d.* AWGN at the receivers are denoted by $W_M \sim \mathcal{CN}(0,1)$, $W_E \sim \mathcal{CN}(0,1)$ for D and E, respectively. Using index i to differentiate one coherence interval from another, the signal received by D and E during one coherence interval are respectively given by

$$\begin{aligned} y(i) =& g_M(i)x(i) + w_M(i), \\ z(i) =& g_E(i)x(i) + w_E(i), \end{aligned} \tag{2.14}$$

where $x(i) \sim \mathcal{CN}(0,1)$ denotes the transmit signal representing a codeword x^n that is mapped from the source message $w \in \mathcal{W} = 1, 2, \ldots, M$ by an (M, n) encoder-decoder. Both the main channel and the eavesdropper channel are assumed to follow the block fading, where $g_M(i)$ and $g_E(i)$ remain constant within each block and vary from block to block. The fading process is also assumed to be stationary and ergodic with bounded continuous probability distribution function (pdf) denoted by $f(h_M)$, $f(h_E)$, for $h_M(i)$ and $h_E(i)$, respectively. In addition, $h_M(i)$ is assumed to be independent of $h_E(i)$, $\forall i$. Denoting channels in one instance by $\underline{h} = (h_M, h_E)$ (dropping the index

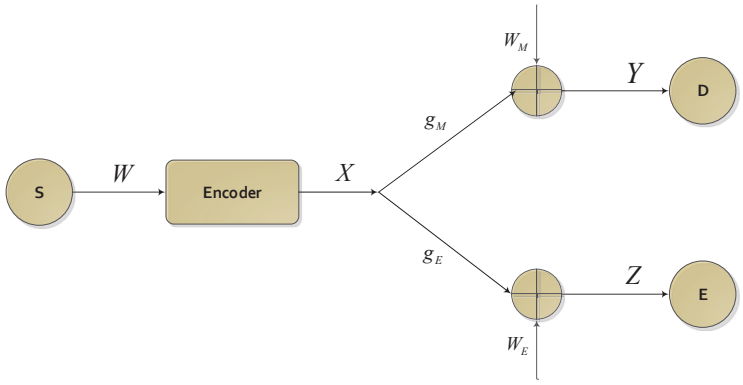


Figure 2.4: A three-user fading wiretap channel.

variable for simplicity), the ESC for the above three-user fading wiretap channel, subject to a long-term power constraint, i.e., $E[p(\underline{h})] \leq P$, is given by [38, 39]

$$C_s = \underset{E[p(\underline{h})]\leq P}{E}[\log_2(1 + h_M p(\underline{h})) - \log_2(1 + h_E p(\underline{h}))], \tag{2.15}$$

where the expectation is taken with respect to (w.r.t.) the RV $\underline{h} \in \mathcal{H}$, with $\mathcal{H} = \{\underline{h}|h_M > h_E\}$. On the other hand, the secrecy outage probability is defined as follows.

$$P_{\text{out}} = \mathrm{P_r}(R_s(\underline{h}, p(\underline{h})) < R_1). \tag{2.16}$$

In (2.16), $R_s(\underline{h}, p(\underline{h}))$ denotes the secrecy capacity for a given fading realization $\underline{h}$ with the source transmit x^n using power $p(\underline{h})$, which is given by [S.K.Leung-Yan-Cheong1978]

$$R_s(\underline{h}, p(\underline{h})) = [\log_2(1 + h_M p(\underline{h})) - \log_2(1 + h_E p(\underline{h}))]^+. \tag{2.17}$$

In the following, some important results on the optimal power allocations to maximize the ESC or minimize the secrecy outage probability are reviewed under the assumption of full and/or partial CSIT, respectively.

1) ESC

- Full CSIT

In this case, the Tx knows the CSI of both the legitimate Rx and the eavesdropper perfectly at the beginning of each coherence interval and therefore is able to adapt the transmit power to the realization of h_M and h_E. Hence, the ESC maximization problem is formulated as

$$\begin{aligned} \text{(P1-full)} \ \underset{\{p(\underline{h})\}}{\texttt{max}} \ & E[\log_2(1 + h_M p(\underline{h})) - \log_2(1 + h_E p(\underline{h}))] \\ \texttt{s.t.} \ \ & E[p(\underline{h})] \leq P, \end{aligned}$$

the expectation of which is still taken w.r.t. $\underline{h} \in \mathcal{H}$. The optimal power allocation policy to (P1-full) is thus given by

$$p(\underline{h}) = \frac{1}{2}\left[\sqrt{\left(\frac{1}{h_E}-\frac{1}{h_M}\right)^2 + \frac{4}{\lambda}\left(\frac{1}{h_E}-\frac{1}{h_M}\right)} - \left(\frac{1}{h_E}+\frac{1}{h_M}\right)\right]^+, \tag{2.18}$$

where the parameter λ is a Lagrangian multiplier that satisfies the long-term power constraint with equality.

It is worthy of noting that (2.18) is sometimes known as the "secrecy water filling" solutions equivalent to those for the SISO fading channel under the long-term power constraint w/o secrecy consideration.

- Partial CSIT

When only the CSIT regarding the legitimate Rx, i.e., h_M is known, the ESC defined in (2.15) is modified as

$$C_s = E[[\log_2(1+h_M p(h_M)) - \log_2(1+h_E p(h_M))]^+]. \tag{2.19}$$

Consequently, the corresponding ESC maximization problem is given by

$$\begin{aligned}(\text{P1-partial}): \max_{\{p(h_M)\}} \; & E[[\log_2(1+h_M p(h_M)) - \log_2(1+h_E p(h_M))]^+] \\ \texttt{s.t.} \quad & E[p(h_M)] \le P,\end{aligned}$$

whose solution, thanks to its convexity, satisfies the following optimality condition [38]:

$$\frac{\partial C_s}{\partial p(h_M)} = \frac{h_M \mathrm{P_r}(h_E \le h_M)}{1+h_M p(h_M)} - \int_0^{h_M}\left(\frac{h_E}{1+h_E p(h_M)}\right) f(h_E)\mathrm{d}h_E - \lambda = 0, \tag{2.20}$$

where λ is a constant that satisfies the long-term power constraint. It is seen that the optimal power allocation policy is also determined by $f(h_E)$, which is known as the channel distribution information (CDI). Hence, the main channel's CSI and at

least the eavesdropper's CDI are assumed to be known at the Tx for this scheme.

Note that compared with (2.15), C_s in the partial CSIT case, the instantaneous secrecy capacity admits $[\cdot]^+$. This subtlety is a consequence of *variable-rate* transmission scheme, in which the Tx transmits at a rate $\log_2(1 + h_M p(h_M))$ adapt to the main channel's fading state. This scheme ensures that when $h_M < h_E$, the achievable rate at the eavesdropper is bounded by $\log_2(1 + h_M p(h_M))$ and when $h_M > h_E$, it is $\log_2(1 + h_E p(h_M))$, which ensures non-negative instantaneous secrecy capacity for the partial CSIT.

2) Secrecy outage probability

- Full CSIT

In this case, the secrecy outage probability minimization problem is as follows.

$$\begin{aligned} \text{(P2-full)}: \min_{\{p(\underline{h})\}} \quad & (2.16) \\ \texttt{s.t.} \quad & E[p(\underline{h})] \leq P, \end{aligned}$$

where $R_1 > 0$ denotes a target secrecy rate. The power allocations policy $p^*(\underline{h})$ that solve (P2-full) for a given target rate R_1 is given by [39, *Proposition 1*]

$$p^*(\underline{h}) = \begin{cases} p^{\min}(\underline{h}), & \text{if } p^{\min}(\underline{h}) \leq s^* \\ 0, & \text{otherwise,} \end{cases} \tag{2.21}$$

where $s^* = \sup\{s | E_{p^{\min}(\underline{h}) \leq s}[p^{\min}(\underline{h})] < P\}$ indicates a threshold solution and $p^{\min}(\underline{h})$ is given by

$$p^{\min}(\underline{h}) = \begin{cases} \frac{2^{R_1}-1}{h_M - 2^{R_1} h_E}, & \text{if } R_1 < \log_2 \frac{h_M}{h_E} \\ \infty, & \text{otherwise.} \end{cases} \tag{2.22}$$

(2.21) is also known as the "secrecy channel inversion" solutions, which are counterparts of those to the outage probability minimization of delay-limited SISO fading channels.

- Partial CSIT

Except for the stochastic properties induced by the fading process, when only partial CSIT is available, the calculation of P_{out} also relies on how the channel uncertainties are modeled [40]. If Alice has full CSI regarding the main channel only the CDI regarding the eavesdropper's as the partial CSIT case discussed above for the ESC, following the similar analysis developed in [38], the optimal power allocations are as follows.

$$p^*(h_M) = \begin{cases} p^{\min}(h_M), & \text{if } \mathcal{L}(\lambda, p^{\min}(h_M)) < 1 \\ 0, & \text{otherwise,} \end{cases} \tag{2.23}$$

where $\mathcal{L}(\lambda, p(h_M)) = \int_{(1/2^{R_1}-1)/p+h_M/2^{R_1}}^{\infty} f(h_E)\mathrm{d}h_E + \lambda p(h_M)$, $p^{\min}(h_M) = \arg\min_{p>(2^{R_1}-1)/h_M} \mathcal{L}(\lambda, p(h_M))$, and λ is the associated optimal dual variable.

For another example, when Alice's knowledge of the eavesdropper's fading coefficient is a noisy version modeled by

$$\hat{g}_E = g_E + w'_E, \tag{2.24}$$

where $\hat{g}_E$ denotes an imperfect estimate of Eve's channel by Alice and $w'_E \sim \mathcal{CN}(0, \sigma_e^2)$ is the stochastic error, the frequently used metric in this situation is also the secrecy outage probability [6]. It admits the same form as (2.16) but has a different operational interpretation, that is, the probability that the wiretap channel fails to support a target secrecy rate R_1. Note that, $R_1 = \log_2(1 + h_M p(h_M)) - \log_2(1 + \hat{h}_E p(h_M))$, where $\hat{h}_E = |\hat{g}_E|^2$ is set by Alice. As a result, perfect secrecy is ensured if Eve's channel is worse than Alice's estimate. However, if Eve's channel is underestimated, i.e., $\hat{h}_E < h_E$, R_1 cannot be achieved and thus secrecy outage occurs. Combing with the fading process, P_{out} under the long-term power constraint $E[p(h_M)] \leq P$ turns out to be as follows.

$$P_{\text{out}} = \int_0^{\infty} \mathrm{P_r}(R_s(h_M, p(h_M)) < R_1 | h_M) f(h_M) \mathrm{d}h_M. \tag{2.25}$$

Although, to the best knowledge of the author, there are no results directly providing optimal power allocations for minimizing P_{out} defined in (2.25), optimal power allocations over all fading realizations are expected to be obtained by similar methods as applied to (P2-full), if $p^{\min}(h_M)$, i.e. the minimum power needed to support the target rate R_1 under the Eve's channel uncertainty, is well approximated. Fortunately, this power minimization problem under the chance constraint w.r.t. channel uncertainties has been well addressed by either equivalently solving a simplified problem in the single-eavesdropper case [32], or conservatively transforming $\mathrm{P_r}(R_s(h_M, p(h_M)) < R_1|h_M)$ into deterministic and convex constraints using *Bernstein-Type Inequality* [40, *Lemma 1*, *Lemma 2*] [41] in the multi-eavesdropper case.

Furthermore, [39] extended the results in [38] to fading broadcast channel with confidential messages (BCC), where the secrecy capacity region for the fading BCC was established, based on which the optimal power allocations were derived to achieve the boundary of the secrecy capacity region, or to minimize the secrecy outage probability with and w/o the common messages, respectively. In addition, [42] employed a slightly different definition of the outage event and combined cryptography and PLS to tackle the secrecy outage in fading channels.

2.2.3 Relay Wiretap Channels

Soliciting for cooperative relaying to improve PLS has drawn much attention since it was first considered in [43], where several cooperation strategies were devised and the corresponding achievable secrecy regions were characterized. Of particular interest was the proposed *noise forwarding (NF)* scheme, where the relay sends codewords independent of source messages while being totally ignorant of the confidential information, which has been shown to increase the secrecy region in the reversely degraded scenario. On the contrary, the conventional relay for cooperative communications fails to provide performance gains in such setup. Benefitting not only from the conventional advantage of cooperative communications, but also from

the effect of CJ on interfering with the eavesdroppers, various relay-assisted secure transmission schemes have thus been investigated, which, depending on the role of the relays, mainly fall into three categories as follows.

1) CB alone

Cooperative beamforming (CB) inherits its concept in the conventional cooperative communications, where relays are exploited to facilitate the source-destination transmission utilising CB via either decode-and-forward (DF) or amplify-and-forward (AF) mode. However, compared with CB in classical relay systems, CB for secrecy takes not only reliable but also secure information transmission into account and therefore yields different optimal beamformer from that w/o secrecy consideration in general. For example, for a single-antenna relay network, the celebrated matched-filter (MF) relay weights, i.e., joint maximum ratio combining (MRC) and maximum ratio transmission (MRT) at the relays, is known to be optimal beamformer for maximizing the communications rate, which is nevertheless not necessarily optimal in terms of maximizing the secrecy rate, since the former solutions may induce large interception at the eavesdroppers as well. As a result, the optimal relay beamformer for maximizing the secrecy rate of the single-antenna DF relay network has been developed in [35] and [44] subject to the total and per-relay power constraints, respectively. For the AF alone case, although the optimal relay beamformer for secrecy rate maximization under a total power constraint was derived in [45], to obtain the optimal relay beamformer under individual relay power constraints poses much challenges and therefore only suboptimal designs are available. [46] extended the work to a multi-eavesdropper scenario under both total and individual relay power constraints, where suboptimal but tractable SDR approaches were exploited. [47] and [48] considered the similar secrecy rate maximization problems for a multi-antenna AF relay network, the former of which developed alternating optimization algorithms to iteratively derive the Tx and relay precoders, while the latter of which studied robust designs for suboptimal beamformers against fading and deterministic channel uncertainty models, respectively.

2) CJ alone

CJ alone refers to the strategy where there is an external relay serving as helper to facilitate the secrecy information transmission, which in particular has two kinds of implementation, i.e., coordinated CJ and uncoordinated CJ (also known as *independent jamming (IJ)*). In coordinated CJ, common jamming signals are cooperatively generated across all single-antenna relays [8, 21, 35, 36, 49], while in uncoordinated CJ, each relay helper emits its own artificial noise (AN) that is independent of the confidential information to confound the eavesdroppers [50, 51].

For the purpose of illustration, a general CJ model that mathematically subsumes all the cases discussed above is assumed in a single-antenna N-relay network, where the CJ signal is given by $\boldsymbol{x} = [x_1, \ldots, x_N]^T$ with its covariance matrix denoted by $\boldsymbol{S} = \tilde{\boldsymbol{V}}\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{V}}^H$, where x_i's are the individual signal transmitted by relays, $\tilde{\boldsymbol{\Sigma}} = diag([\sigma_1, \ldots, \sigma_d])$ is a diagonal matrix comprising all the positive eigenvalues of $\boldsymbol{S}$, and $\tilde{\boldsymbol{V}}$ is the associated precoding matrix satisfying $\tilde{\boldsymbol{V}}^H\tilde{\boldsymbol{V}} = \boldsymbol{I}$. Accordingly, the generation of $\boldsymbol{x}$ is uniquely determined by $\boldsymbol{S}$ in the sense that $\boldsymbol{x} = \sum_{j=1}^{d}\sqrt{\sigma_j}\tilde{\boldsymbol{v}}_j s'_j$, where $\tilde{\boldsymbol{v}}_j$'s are drawn from the columns of $\tilde{\boldsymbol{V}}$ and $s'_j \sim \mathcal{CN}(0,1)$, $\forall j$. For a special case of $d = 1$, i.e., $\boldsymbol{x}_r = \sqrt{\sigma_1}\tilde{\boldsymbol{v}}_1 s'_1$, each relay transmits one common jamming signal s'_1 with their respective weight drawn from $\tilde{\boldsymbol{v}}_1$. This kind of coordinated CJ is desirable in practice since it reduces the overhead of exchanging d CJ beams. Another special case when $\boldsymbol{S}$ is a full-rank diagonal matrix with all σ_i's positive is referred as IJ, which corresponds to the jamming scheme where each relay transmits *i.i.d.* CJ beam s'_i, $i = 1, \ldots, N$. It is worthy of noting that IJ is of practical interest, since the jamming beams s'_i's can be completely generated in a distributed fashion without coordination among relays.

3) Joint CB and CJ

Under the circumstances that the direct link between the Tx and the legitimate Rx is broken, i.e., some of the relays have to take on their conventional role of forwarding the information apart from jamming. Depending on whether information forwarding

and CJ are performed on the same relay, the joint CB and CJ schemes vary in their designs. For separately located conventional relays and friendly jammers, a selection of their function was studied in [52–54], while for co-located relay helpers, a recent paradigm makes better use of d.o.f of relays by allowing simultaneously forwarding the confidential information and emitting AN [55, 56].

The CB and AN design for secrecy rate maximization from a single-antenna Tx to a single-antenna Tx via multiple multi-antenna relays in the presence of multiple multi-antenna eavesdroppers, in spite of being more general, is quite challenging to solve due to the amplified noise at relays, and nonconvex and nonsmooth property of the secrecy rate function etc. [51] computed the joint optimal CB and AN solution to the above problem under imperfect CSIT w.r.t. the eavesdroppers' channels using SDR, which was proved be tight. For a more practical scenario when no Eve's CSI is available, similar to [30], in [31], distributed beamforming with AN was performed by associated intermediate nodes so as to maximize the power permitted for AN under individual power constraint of each AF relay. It has been noticed that even for the case of perfect CSIT, the joint transmission beams, relay beams and CJ optimization for secrecy MIMO relay networks turns out to be very intractable and only some simplified designs are available in the literature. For example, interference alignment-based CJ was employed in [56] for a single multi-antenna relay MIMO wiretap channel.

2.3 Recent Progress in Enhancing PLS for SWIPT Systems

Some major challenges and opportunities identified in the literature for wireless PLS in WPCN are surveyed in this section mainly from a signal processing point of view, which fundamentally differ these SWIPT-enabled PLS enhancements from those reviewed in Section 2.2. The main thrust of this thesis is also developed based on and/or in parallel with some of the results reviewed in this section. Complying

with the motivation introduced in Chapter 1, the challenges of enhancing PLS in WPCN are introduced first followed by opportunities in related work.

[57] and [58] were among the earliest work that identified the critical threat to PLS in SWIPT, that is the confidential messages destined to IRs may be eavesdropped by ERs who need to operate with significantly higher received power as compared to the conventional IRs and are therefore usually deployed in more proximity to the Tx than the IRs. Both of the work advocated the usage of AN and/or energy signal to protect the secrecy information against eavesdropping in a MISO downlink system with one IR and multiple ERs when different levels of eavesdroppers' CSIT is available. Specifically, joint design of transmit and AN beamforming vectors along with their power allocations was investigated in [57] with different objectives: the first problem maximized the secrecy rate for the IR subject to individual EH constraints of ERs; the second problem maximized the weighted sum-energy transferred to ERs subject to the required secrecy rate for the IR. Both of these problems were non-convex and optimally solved by a two-stage procedure utilizing the technique of SDR. Furthermore, two suboptimal solutions of lower complexity that separately design the information and AN beamforming vectors were proposed for each of the studied problems, and then compared against the optimal solution in terms of achievable (secrecy) rate-energy trade-off. While [57] considered the secure beamforming schemes with perfect eavesdroppers' CSI known at the Tx, [58] investigated a more complex scenario where the CSI of the desired IR is perfectly available whereas that of the idle receivers (potential eavesdroppers) and passive eavesdroppers are only imperfectly and not known at the Tx, respectively. Compared with [57], the non-convexity was also incurred by the minimum outage probability requirement at the passive eavesdroppers, which was then replaced by a convex deterministic constraint. The advantage of the joint design of AN and/or energy signals as well as the information beamforming further corroborated the dual usage of AN that facilitates secrecy SWIPT, which is fundamentally different from the mere role of interfering with eavesdroppers in conventional wiretap channels.

[59] further considered the secure beamforming designs in a more threatening scenario where the ERs are able to collude each other to perform joint decoding of the confidential messages, and therein obtained the robust beamforming against the eavesdropper's channel uncertainties including both covariance-based and worst-cased based imperfect CSI, respectively. A similar worst-case robust secrecy rate maximization problem was also considered in [60] to cripple multiple multi-antenna ERs's eavesdropping.

[61] considered a more general system model that enables secure multi-casting for SWIPT (with multiple IRs) in the presence of multiple ERs acting as potential eavesdroppers. Unlike designing the optimal transmit beamforming based on SDR, the authors proposed a novel secure multi-casting design to minimize the total transmit power under the SINR constraints for IRs/ERs and the harvested energy constraints for ERs by using rank-two beamformed Alamouti space-time (AST) coding and SDR, apart from jointly employing energy signals as AN. The sufficient conditions under which the proposed scheme is optimal were also derived with a corresponding rank-two Gaussian randomization procedure, which provides a suboptimal solution when the SDR is not tight. [62] extended the system model to a three-node MIMO downlink system with one IR and one eavesdropping ER. The authors aimed to maximize the achievable secrecy rate subject to a transmit power constraint and an EH constraint for the ER. The SRM problem with multiple streams was then solved by an inexact block coordinate descent (IBCD) algorithm, which proved to monotonically converge to a Karush-Kuhn-Tucker (KKT) solution. Besides, [63] explored secure beamforming design in a multi-user MISO secondary communication system with secondary idle receivers as potential eavesdroppers using multi-objective optimization, while [64] studied the secrecy precoding for a MIMOME SWIPT-enabled secondary communication system. In addition, a massive MIMOME wiretap channel for a SWIPT system was studied in [65], which derived the asymptotic-optimal transmit covariance that achieves the trade-off between the ESC and the harvested energy given only statistical CSIT using large-dimension random

matrix theory and Tayler series expansion.

Following the trend of secure beamforming design in SWIPT, in contrast with the aforementioned work that considered co-located ERs and eavesdroppers, [66] and [67] studied the secure information transmission with separate ERs and eavesdroppers (passive) under perfect and imperfect eavesdroppers' CSI, respectively. The secrecy optimization framework with and w/o AN for MISO wiretap channel was also investigated in [68] by taking error-bounded channel uncertainties into account. Moreover, [68] proposed successive convex approximation-based solutions to the associated problems as effective alternative for those rank-relaxation based approaches in the literature [69]. Later, when an external help was solicited for assisting the source in secrecy SWIPT with EH constraints for both IR and ERs, a CJ-aided worst-case secrecy rate maximization problem was considered in [70] by employing alternating optimization.

On another front, with the upsurge of SWIPT, when some of the idle WEH-enabled network subscribers are friendly and not subject to stringent EH requirements, recent advances in WPCN [71] provide essential incentives for them to assist in the secrecy communications. [72] considered to boost the confidential information transmission throughput (long-term metric) via a wireless-powered friendly jammer by proposing a judiciously designed protocol including the "harvest" and "jam" phases, based on which four types of power transfer (PT)-information transmission (IT) cycles were characterized and the behavior of this stochastic process was analyzed to derive a closed-form throughput. In addition, the jamming power threshold and rate parameters were further optimized to maximize the throughput subject to a secrecy outage probability constraint. Furthermore, motivated by the PS technologies which circumvent the circuit limit that the received signal used for harvesting energy cannot be reused for decoding information, wireless-powered relays adopting the PS-enabled hybrid receiver have recently been exploited to enhance PLS by means of self-sustaining cooperative strategies [73, 74]. Specifically, [73] investigated the joint optimization of PS ratios and relay weights to maximize the

secrecy rate in a multiple non-regenerative wireless-powered relay networks resorting to destination-based AN. The authors proposed efficient numerical algorithm that converges to a local optimal solution. A similar approach was further applied to an untrustful multi-antenna wireless-powered relay network by [74]. Alike [73], the WPCN was in the favor of PLS design in the sense that the destination-generated AN also served as a new source of RF EH. Very recently, motivated by the 5G-enabling large-scale MIMO technology, secret transmission assisted by a wireless-powered large-scale MIMO relay was considered in [75], where an explicit expression of the secrecy outage capacity was derived under the assumption of imperfect legitimate Rx's CSI and no eavesdropper's CSI.

Chapter 3

Secrecy Wireless Information and Power Transfer in Fading Wiretap Channel

3.1 Introduction

SWIPT has recently drawn significant interests for its dual use of radio signals to provide wireless data and energy access at the same time. In a SWIPT system with secrecy information transmission to the IRs, there are two-fold conflicting goals in the transmission design: the distance of ERs away from the hybrid AP is scheduled short enough to circumvent the low energy efficiency of WPT due to the substantial power attenuation of RF signals, but is also expected to be longer than that of IRs in order for achieving *perfect secrecy* if they do not harvest energy as presumed; the power of the received signal at the ERs is desired to be made large for efficient WEH, but also needs to be kept sufficiently small to prevent leakage of the borne confidential information. To resolve this conflict, in this chapter the transmit signal is split into two parts, with one part carrying the secrecy information for the IR and the other part carrying an AN to interfere with the ER to prevent it from eavesdropping, while the total signal power received at the ER can still be kept high to satisfy its energy harvesting requirement, under the assumption that the AN can be cancelled at IRs but not at ERs. Under a simplified three-node wiretap channel setup, the transmit power allocations and power splitting ratios over fading channels are jointly optimized to minimize the outage probability for delay-limited secrecy information transmission, or to maximize the average rate for no-delay-limited secrecy information transmission, subject to a combination of average and peak power constraints at the transmitter

as well as an average energy harvesting constraint at the ER. For each of these non-convex problems the optimal solution based on the dual decomposition as well as suboptimal solution based on the alternating optimization are proposed. Finally, the performances of proposed schemes are evaluated by simulations against two benchmark schemes in terms of various trade-offs for wireless (secrecy) information versus power transmissions.

3.1.1 Assumption

Note that unlike the existing literature on PLS, where the eavesdroppers are passive devices and thus their channels are practically assumed to be unknown at the Tx, in this chapter, the Tx is assumed to know the ER's eavesdropping channel since the ER needs to assist the Tx in obtaining its CSI to design the power allocations to satisfy its energy harvesting requirement. Moreover, for the AN-aided transmission, both the Tx and IR are assumed to have the knowledge of the AN to be used prior to transmission via a known PHY-layer "key" distribution method [76, 77] (see Section 3.3 for the details). Thus, the AN can be cancelled at the IR and however, is kept strictly confidential to the ER so that it cannot be cancelled at the ER. Such a scheme provides a theoretical upper-bound for the achievable secrecy rate of the SWIPT system under our consideration.

The remainder of the chapter is organized as follows. The SWIPT system model over a SISO fading wiretap channel is introduced in Section 3.3. The formulations of the proposed secrecy outage probability minimization problem and the ESC maximization problem are presented in Section 3.4. Both optimal and suboptimal solutions to the two formulated problems are proposed in Section 3.5 and Section 3.6, respectively. Two benchmark schemes and their optimal designs are presented in Section 3.7. Numerical results on the performance of various schemes proposed are provided in Section 3.8. Finally, the chapter is concluded in Section 3.9.

3.2 Related Work

3.2.1 AN in Secrecy SWIPT

In the conventional secrecy communication setup without the EH consideration, AN has been widely applied to improve the secrecy transmission rates [21, 33, 34, 78], where a fraction of the transmit power was allocated to send randomly generated noise signals to reduce the amount of information intercepted by the eavesdroppers. In [57], AN was first applied in a MISO SWIPT system, where the joint information and energy beamforming design at the transmitter was investigated to maximize the secrecy rate of the IR subject to individual harvested energy constraints of ERs, or to maximize the weighted sum-power harvested by ERs subject to a given secrecy rate constraint at the IR. However, [57] considered the AWGN channels, while the optimal AN-aided secrecy transmission design for SWIPT systems over fading channels has not yet been addressed in the literature, which motivates the work in this chapter.

3.2.2 The Role of Fading in PLS

Although channel fading is traditionally regarded as a detrimental factor to the wireless channel capacity, it can be exploited to reduce the secrecy communication outage probability [6, 37, 39, 42, 79] or improve the wireless channel secrecy capacity [6, 38, 39, 80]. For the secrecy outage probability minimization for wireless fading channels with stringent transmission delay constraint, [39] has derived the optimal power allocations in the fading broadcast channel with confidential messages assuming the CSIT is known. While for maximizing the ESC of fading channels with no-delay-limited transmission, the corresponding optimal power and rate allocation strategies have been studied in [38]. However, existing results for fading wiretap channels cannot be directly applied in our new SWIPT setup due to the additional energy harvesting requirement for the ER (which may also play a role of eavesdropper).

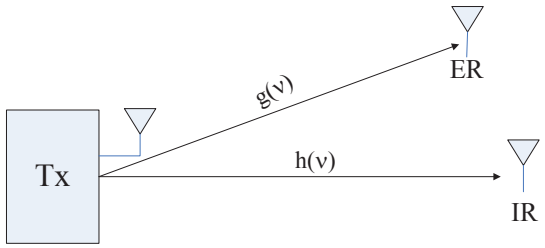


Figure 3.1: The fading wiretap channel in a three-node SWIPT system.

3.3 System Model

In this chapter, consider a three-node SISO fading wiretap channel in a SWIPT system consisting of one transmitter (Tx), one IR and one ER, each equipped with one antenna, as shown in Fig. 3.1. The complex channel coefficients from the Tx to IR and ER for one particular fading state are denoted by $u(\nu)$ and $v(\nu)$, respectively, where ν denotes the joint fading state. The power gains of the channels at fading state ν are defined as $h(\nu) = |u(\nu)|^2$ and $g(\nu) = |v(\nu)|^2$; and it is assumed that at each fading state ν, both $h(\nu)$ and $g(\nu)$ are perfectly known at the Tx. [1] A block fading model is further assumed such that $h(\nu)$ and $g(\nu)$ remain constant during each block for each fading state ν, but can vary from block to block as ν changes. It is assumed that $h(\nu)$ and $g(\nu)$ are two RVs with a continuous joint pdf.

Since secrecy information transmission to the IR is of the interest, similar to [21], the transmit signal is assumed to comprise an information-bearing signal s_0 and an AN-bearing signal s_1. It is assumed that s_0 is a circularly symmetric complex Gaussian (CSCG) RV with zero mean and unit variance, denoted by $s_0 \sim \mathcal{CN}(0, 1)$. Furthermore, since s_1 plays the role of AN to reduce the information eavesdropped by the ER and the worst case AN is known to be Gaussian distributed [21], s_1 is also

[1] In practice, considering time division duplex (TDD) is used, at the beginning of each transmission block, the IR and ER can send their respective pilot signal to the Tx for it to estimate the reverse-link channel assuming short-term channel reciprocity between the Tx and IR/ER. TDD is also assumed for the subsequent description of secret "key" generation and transmission.

assumed to be a CSCG RV denoted by $s_1 \sim \mathcal{CN}(0,1)$, and independent of s_0. The complex baseband transmit signal at fading state ν is thus expressed as

$$x = \sqrt{(1-\alpha(\nu))p(\nu)}s_0 + \sqrt{\alpha(\nu)p(\nu)}s_1, \tag{3.1}$$

where $p(\nu)$ is the transmit power at fading state ν and $0 \leq \alpha(\nu) \leq 1$ denotes the portion of the transmit power allocated to the AN signal at fading state ν. Moreover, similar to [81], in this chapter two types of power constraints on $p(\nu)$ are considered, namely, average power constraint (APC) and peak power constraint (PPC). The APC limits the average transmit power at the Tx over all fading states, i.e., $E_\nu[p(\nu)] \leq P_{\text{avg}}$, where $E_\nu[\cdot]$ denotes the expectation over ν. In contrast, the PPC constrains the instantaneous transmit power of the Tx at each fading state, i.e., $p(\nu) \leq P_{\text{peak}}$, $\forall \nu$. W.l.o.g., $P_{\text{avg}} \leq P_{\text{peak}}$ is assumed. The signals received at the IR and the ER are then respectively given by

$$\begin{aligned} y_{\text{IR}} &= u(\nu)x + n_{\text{IR}} \\ &= u(\nu)\left(\sqrt{(1-\alpha(\nu))p(\nu)}s_0 + \sqrt{\alpha(\nu)p(\nu)}s_1\right) + n_{\text{IR}}, \end{aligned} \tag{3.2}$$

$$\begin{aligned} y_{\text{ER}} &= v(\nu)x + n_{\text{ER}} \\ &= v(\nu)\left(\sqrt{(1-\alpha(\nu))p(\nu)}s_0 + \sqrt{\alpha(\nu)p(\nu)}s_1\right) + n_{\text{ER}}, \end{aligned} \tag{3.3}$$

where $n_{\text{IR}} \sim \mathcal{CN}(0,\sigma_1^2)$ and $n_{\text{ER}} \sim \mathcal{CN}(0,\sigma_2^2)$ denote the AWGN at the IR and the ER, respectively.

3.3.1 A PHY-layer "key" Distribution Scheme

As previously mentioned in the chapter, a PHY-layer "key" distribution scheme with practical complexity is assumed for generating and cancelling the AN signal s_1 at the IR, which is described in this subsection. First, a large ensemble of seeds for a Gaussian pseudo-random generator are pre-stored at both the Tx and IR (but not available at the ER). The index of each seed in the ensemble is denoted as a "key" in

the sequel. Next, by randomly picking up one seed and transmitting its index to the IR before sending the confidential message at the beginning of each fading state, the Tx is able to generate a "random" AN sequence using the selected seed that is only known to the IR. Note that the seed used at each fading state is random and unknown to the ER since "key" (index of the seed in use) is also non-accessible by the ER. To achieve such secure "key" sharing, a two-step phase-shift modulation based method [76, 77] by leveraging the short-term reciprocity of the wireless channels between the Tx and IR is further adopted.

Specifically, in the first step, IR transmits a sinusoid pilot at frequency f_c to the Tx in the kth interval $(kT, (k+1)T]$, for example, which is given by

$$s_{\mathrm{IR}}(t) = \sqrt{\frac{2E}{T}} \cos(2\pi f_c t + \phi), \tag{3.4}$$

where E is the symbol energy and ϕ is a reference phase. The received signal by the Tx as a result, given by

$$r_{\mathrm{Tx}}(t) = \sqrt{\frac{2\Lambda^2(k)E}{T}} \cos(2\pi f_c t + \Theta(k)) + n(t), \tag{3.5}$$

where $\Lambda^2(k)$'s denote the power gain due to the large-scale fading and remain the same as $h(\nu)$ for each fading state, while $\Theta(k)$'s indicate the small-scale fading that can be differentially estimated by the Tx w.r.t. ϕ as $\Theta(k) - \phi$. $n(t)$ is the received AWGN. Now that the Tx has already probed the channel response from the IR, under the assumption of reciprocity of channels, it will in the sequel use this CSI to transmit a "key" to the IR by pre-compensating the phase difference of the sinusoid in the same frequency as follows.

$$s_{\mathrm{Tx}}(t) = \sqrt{\frac{2E}{T}} \cos(2\pi f_c t - (\Theta(k) - \phi) + \Psi), \tag{3.6}$$

where $\Psi \in \{-\pi, \ldots, -\pi + 2(M-1)\pi/M\}$ is determined by the "key"-bearing symbol and the corresponding constellation for modulation and M indicates cardinality of

the "key" symbol set, i.e., the number of phase decoding regions. As a result, IR receives

$$r_{\mathrm{IR}}(t) = \sqrt{\frac{2\Lambda^2(k)E}{T}}\cos(2\pi f_c t + \phi + \Psi) + n(t), \tag{3.7}$$

which allows the IR to decode the information borne in the estimated phase difference, i.e., Ψ, in accordance with the region that it falls in.

It is worth noting that the effect of the proposed "key" sharing closely depends on two assumptions: the channels' reciprocity and the de-correlation of phase differences. The former is ensured as long as the mobile's distance moved within the transmission interval is negligible compared to the wavelength. The latter is almost satisfied since the difference between the estimated phase at the ER and that at the Tx has been shown to be almost uniformly distributed [77], which thus enforces the ER to break the "key" only by exhaustive trials.

Remark 3.3.1. *The reason for which the above scheme is not employed to transmit information is as follows. Since the sequence of symbols are transmitted within a block duration so that the symbol interval T is less than the channel coherence time. Thus, the received phase difference between the symbols being transmitted in two successive intervals seen by ER would fall in the information set $\{-\pi, \ldots, -\pi + 2(M-1)\pi/M\}$, which is prohibitive in terms of "strong secrecy". However, it is relatively secure for the "key" transmission, since even if the ER attempts to decode the seed index, it does not have access to the seed ensemble, the complexity for breaking which is practically infeasible.*

Remark 3.3.2. *If the Tx and the IR are assumed to share certain common information a priori, our considered scheme may not be optimal as inspired by [42, 79]. Nevertheless, this scheme is considered for its ease of implementation in view of two folds. For one thing, in practical SWIPT systems the AN also plays the role of delivering wireless power to the ER. For the other thing, although the above "key" distribution method requires additional transmission time, it is negligible*

compared to the whole length of block duration especially when the channel coherence time is sufficiently large.

With the scheme proposed in the last subsection, the associated interference at the IR in (3.2), i.e., $u(\nu)\sqrt{\alpha(\nu)p(\nu)}s_1$, can be canceled at each fading state prior to decoding the desired information signal, s_0. Then from (3.2), the SNR at the IR at fading state ν with a given pair of $\alpha(\nu)$ and $p(\nu)$ is expressed as

$$\mathrm{SNR}_{\mathrm{IR}}(\alpha(\nu), p(\nu)) = \frac{(1-\alpha(\nu))h(\nu)p(\nu)}{\sigma_1^2}. \tag{3.8}$$

Note that in practice the AN cancelation at the IR cannot be perfect, while the residue interference due to imperfect AN cancellation could be included in the receiver noise power, i.e., σ_1^2. On the other hand, since the AN signal s_1 is assumed to be unknown to the ER and thus cannot be canceled, from (3.3), the SNR at the ER at fading state ν is expressed as (assume that the ER eavesdrops the information intended for the IR instead of harvesting energy)

$$\mathrm{SNR}_{\mathrm{ER}}(\alpha(\nu), p(\nu)) = \frac{(1-\alpha(\nu))g(\nu)p(\nu)}{\alpha(\nu)g(\nu)p(\nu)+\sigma_2^2}. \tag{3.9}$$

Then, the achievable secrecy rate at fading state ν can be expressed as [21]

$$\begin{aligned} R(\alpha(\nu), p(\nu)) = &\Bigg[\log_2\left(1+\frac{(1-\alpha(\nu))h(\nu)p(\nu)}{\sigma_1^2}\right) \\ &- \log_2\left(1+\frac{(1-\alpha(\nu))g(\nu)p(\nu)}{\alpha(\nu)g(\nu)p(\nu)+\sigma_2^2}\right)\Bigg]^+, \end{aligned} \tag{3.10}$$

where $[x]^+ \triangleq \max(0, x)$.

Next, for wireless power transfer, the amount of power harvested at fading state ν at the ER is given by [82]

$$\begin{aligned} Q(p(\nu)) &= \zeta\left[(1-\alpha(\nu))g(\nu)p(\nu) + \alpha(\nu)g(\nu)p(\nu)\right] \\ &= \zeta g(\nu)p(\nu), \end{aligned} \tag{3.11}$$

where $0 < \zeta \leq 1$ denotes the energy harvesting efficiency. Note that the background noise power σ_2^2 is ignored in (3.11), since it is typically very small as compared with the received signal power for energy harvesting. The average harvested power at the ER is thus given by

$$Q_{\text{avg}} = E_\nu \left[Q(p(\nu)) \right]. \tag{3.12}$$

3.4 Problem Formulation

In this section, both delay-limited and no-delay-limited secrecy information transmission to the IR are considered, for which the design problems are formulated in the following two subsections, respectively.

3.4.1 Delay-Limited Secrecy Information Transmission

First, consider the delay-limited secrecy information transmission to the IR, for which the outage probability is a relevant metric. Given a target rate r_0, the secrecy outage probability at the IR can be expressed as [39]

$$\delta = Pr(R(\alpha(\nu), p(\nu)) < r_0), \tag{3.13}$$

where $R(\alpha(\nu), p(\nu))$ is the achievable secrecy rate at fading state ν given in (3.10), and $Pr(\cdot)$ denotes the probability. With CSIT known, the transmitter-aware secrecy outage probability is generally minimized by the "secrecy channel inversion" based power allocation strategies [39]. For convenience, the following indicator function is introduced for the event of outage w.r.t. the target secrecy rate r_0 at each fading state ν:

$$X(\nu) = \begin{cases} 1 & \text{if } R(\alpha(\nu), p(\nu)) < r_0, \\ 0 & \text{otherwise.} \end{cases} \tag{3.14}$$

It thus follows that the outage probability can be re-expressed as $\delta = Pr(R(\alpha(\nu), p(\nu)) < r_0) = E_\nu[X(\nu)]$.

For delay-limited secrecy information transmission, we aim at minimizing the secrecy outage probability for the IR by jointly optimizing the transmit power allocations, i.e., $\{p(\nu)\}$, as well as the transmit power splitting ratios, i.e., $\{\alpha(\nu)\}$ over different fading states, subject to a given pair of combined APC and PPC at the Tx, i.e., P_{avg} and P_{peak}, as well as an average harvested power constraint at the ER, denoted by $\bar{Q}$. Therefore, we consider the following optimization problem.

$$\begin{aligned} (\text{P1}): \ \underset{\{p(\nu),\alpha(\nu)\}}{\texttt{Minimize}} \quad & E_\nu[X(\nu)] \\ \texttt{Subject to} \quad & E_\nu[p(\nu)] \leq P_{\text{avg}}, \\ & p(\nu) \leq P_{\text{peak}}, \ \forall \nu, \\ & E_\nu[Q(p(\nu))] \geq \bar{Q}, \\ & 0 \leq \alpha(\nu) \leq 1, \ \forall \nu. \end{aligned}$$

3.4.2 No-Delay-Limited Secrecy Information Transmission

Next, consider the no-delay-limited secrecy information transmission to the IR. In this case, ESC is a relevant metric that is expressed as

$$C_s = E_\nu[R(\alpha(\nu), p(\nu))]. \tag{3.15}$$

With CSIT known, (3.15) is generally maximized by the "secrecy water-filling" based power allocation policies [38, 39].

For no-delay-limited secrecy information transmission, we aim at maximizing the ESC for the IR subject to the same set of constraints (APC, PPC at the Tx, and an average harvested power constraint at the ER) as for the delay-limited case in (P1).

Therefore, we consider the resulting optimization problem as follows.

$$\begin{aligned}
(\text{P2}): \ \underset{\{p(\nu),\alpha(\nu)\}}{\texttt{Maximize}} \quad & E_\nu[R(\alpha(\nu), p(\nu))] \\
\texttt{Subject to} \quad & E_\nu[p(\nu)] \leq P_{\text{avg}}, \\
& p(\nu) \leq P_{\text{peak}}, \forall \nu, \\
& E_\nu[Q(p(\nu))] \geq \bar{Q}, \\
& 0 \leq \alpha(\nu) \leq 1, \forall \nu.
\end{aligned}$$

Since the objective functions in (P1) and (P2) are in general non-convex and non-concave, respectively, (P1) and (P2) are non-convex problems. In the following two sections, we propose both optimal and suboptimal solutions to these two problems, respectively.

3.5 Proposed Solutions for Delay-Limited Case

In this section, we propose both optimal and suboptimal solutions to (P1). First, we derive the optimal power allocations, i.e., $\{p(\nu)\}$, and power splitting ratios, i.e., $\{\alpha(\nu)\}$, to solve problem (P1). Although (P1) is shown to be non-convex, it can still be solved effectively with global optimum solutions thanks to a so-called "time-sharing" condition proposed in [2], which is introduced shortly. "time-sharing" condition guarantees that zero duality gap approximately holds for (P1) so that the Lagrangian duality method can be applied subsequently.

3.5.1 Time-Sharing Condition

The objective of this section is to interpret that the *time-sharing* condition implies zero duality gap and the problems of our interest such as (P1) and (P2) satisfy this condition under the assumption of continuous joint distribution of $h(\nu)$ and $g(\nu)$. To generalize the context in which this property is usually embedded,

consider a general optimization problem of the following form:

$$(\text{P0}): \quad \max_{\{\boldsymbol{x}_n\}} \quad \sum_{n=1}^{N} f_n(\boldsymbol{x}_n)$$
$$\text{s.t.} \quad \sum_{n=1}^{N} \boldsymbol{h}_n(\boldsymbol{x}_n) \leq \boldsymbol{P},$$

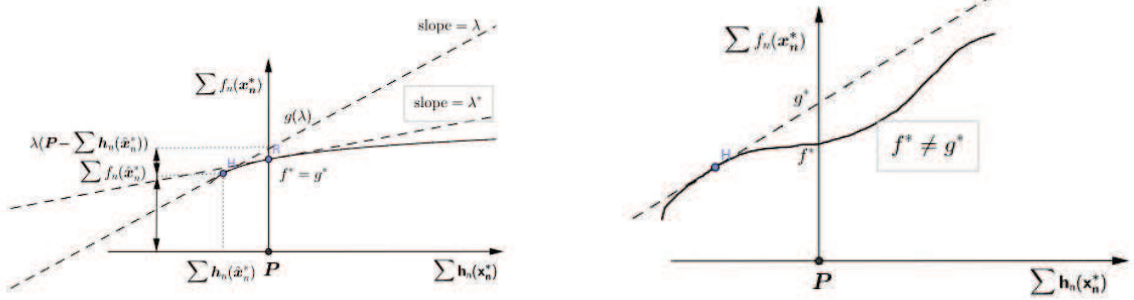
where $\boldsymbol{x}_n \in \mathbb{R}^K$ is a vector of optimization variables, i.e., $\boldsymbol{x}_n = (x_1^n, \ldots, x_K^n)^T$, $\forall n$, $f_n(\cdot)$'s are $\mathbb{R}^K \to \mathbb{R}$ functions, and $\boldsymbol{h}_n(\cdot)$ are functions that map from $\mathbb{R}^K$ to $\mathbb{R}^L$. The L power constraints are denoted by an L-vector $\boldsymbol{P} \in \mathbb{R}^L$ and "$\leq$" stands for component-wise no bigger than.

Definition 3.5.1. *([2,* Definition 1 *]) Let $\boldsymbol{x}_n^*$ and $\boldsymbol{y}_n^*$ be optimal solutions to the optimization problem (P0) with $\boldsymbol{P} = \boldsymbol{P}_x$ and $\boldsymbol{P} = \boldsymbol{P}_y$, respectively. (P0) is said to satisfy the* time-sharing *condition if $\forall \boldsymbol{P}_x$, $\forall \boldsymbol{P}_y$, and $\forall \theta \in [0,1]$, $\exists \boldsymbol{z}_n$'s that are feasible solutions to (P0), such that $\sum_{n=1}^{N} \boldsymbol{h}_n(\boldsymbol{z}_n) \leq \theta \boldsymbol{P}_x + (1-\theta)\boldsymbol{P}_y$, and $\sum_{n=1}^{N} f_n(\boldsymbol{z}_n) \geq \theta \sum_{n=1}^{N} f_n(\boldsymbol{x}_n^*) + (1-\theta) \sum_{n=1}^{N} f_n(\boldsymbol{y}_n^*)$.*

Lemma 3.5.1. *If problem (P0) satisfies the "time-sharing" condition defined in definition 3.5.1, then the optimum value of problem (P0), is a concave function of $\boldsymbol{P}$.*

Proof. Let $\boldsymbol{P}_z$ be a vector with $\boldsymbol{P}_z = \theta \boldsymbol{P}_x + (1-\theta)\boldsymbol{P}_y$ and $\boldsymbol{z}_n^*$ be the optimal solutions to problem (P0) with the constraint $\boldsymbol{P}_z$. For the convenience of exposition, we denote the optimum value of problem (P0) with a constraint $\boldsymbol{P}$ by $f^*(\boldsymbol{P})$. Since for any $0 \leq \theta \leq 1$, there always exists a $\boldsymbol{z}_n$ such that $\sum_{n=1}^{N} \boldsymbol{h}_n(\boldsymbol{z}_n) \leq \theta \boldsymbol{P}_x + (1-\theta)\boldsymbol{P}_y = \boldsymbol{P}_z$, the feasibility implies that $f^*(\boldsymbol{P}_z) = \sum_{n=1}^{N} f_n(\boldsymbol{z}_n^*) \geq \sum_{n=1}^{N} f_n(\boldsymbol{z}_n)$, which is, on the other hand, no smaller than $\theta \sum_{n=1}^{N} f_n(\boldsymbol{x}_n^*) + (1-\theta) \sum_{n=1}^{N} f_n(\boldsymbol{y}_n^*) = \theta f^*(\boldsymbol{P}_x) + (1-\theta) f^*(\boldsymbol{P}_y)$. Combining the above two facts, it follows that $f^*(\boldsymbol{P}_z) = f^*(\theta \boldsymbol{P}_x + (1-\theta)\boldsymbol{P}_y) \geq \theta f^*(\boldsymbol{P}_x) + (1-\theta) f^*(\boldsymbol{P}_y)$, which completes the proof. □

Next, we show that the concavity of $f^*(\boldsymbol{P})$ w.r.t. $\boldsymbol{P}$ implies zero duality gap. Theoretical proof can be referred to [83] while a graphical illustration of



(a) An optimization problem that satisfies the "time-sharing" condition.

(b) An optimization problem w/o "time-sharing" condition.

Figure 3.2: "Time-sharing" condition implies zero duality gap [2].

the proof when $L = 1$ is given in the sequel. As it is easily verified that (P0) achieves its optimum value when the inequality is active, Lemma 3.5.1 suggests that $\sum_{n=1}^{N} f_n(x_n^*)$ is also concave w.r.t. $\sum_{n=1}^{N} h_n(x_n^*)$ as shown by the solid line in Fig. 3.2(a), in which $\sum_{n=1}^{N} h_n(x_n^*)$ and $\sum_{n=1}^{N} f_n(x_n^*)$ denote the coordinates corresponding to the x-axis and y-axis, respectively. Given a fixed P, the intersection of the curve $(\sum_{n=1}^{N} h_n(x_n^*), \sum_{n=1}^{N} f_n(x_n^*))$ with $\sum_{n=1}^{N} h_n(x_n^*) = P$ yields exactly the primal optimum value f^* to problem (P0), as denoted by R in Fig. 3.2(a). Consider the dual objective function $g(\lambda)$ associated with (P0) given by

$$g(\lambda) = \max_{\{x_n\}} \{\sum_{n=1}^{N} f_n(x_n) + \lambda(P - \sum_{n=1}^{N} h_n(x_n))\}. \tag{3.16}$$

Given any λ, let $\hat{x}_n^*$ be the optimal solutions to (3.16). Then $g(\lambda)$ can be graphically illustrated by a line tangent to the point $(\sum_{n=1}^{N} h_n(\hat{x}_n^*), \sum_{n=1}^{N} f_n(\hat{x}_n^*))$ (c.f. H in Fig. 3.2(a)) on the curve $(\sum_{n=1}^{N} h_n(x_n^*), \sum_{n=1}^{N} f_n(x_n^*))$. Consequently, λ is equivalent to the slope of this line, while $g(\lambda) = \sum_{n=1}^{N} f_n(\hat{x}_n^*) + \lambda(P - \sum_{n=1}^{N} h_n(\hat{x}_n^*))$ happens to be the y-coordinate of the intersection of this line with $\sum_{n=1}^{N} h_n(x_n^*) = P$. Now

recall the dual problem of (P0) as follows.

$$\min_{\lambda \geq 0} g(\lambda).$$

The minimization of $g(\lambda)$ can therefore be visualized as the minimum coordinate of the tangent line intersected by $\sum_{n=1}^{N} h_n(x_n^*) = P$, which is achieved by the line tangent to (P, f^*), i.e., $g^* = f^*$, as seen in Fig. 3.2(a). Hence we have proved that the concavity of $f^*(\boldsymbol{P})$ w.r.t. $\boldsymbol{P}$ suffices for a zero duality gap between the primal problem (P0) and its dual. To the end of comparison, Fig. 3.2(b) depicts a situation when "time-sharing" condition does not hold and therefore $\sum_{n=1}^{N} f_n(x_n^*)$ is not necessarily concave w.r.t. $\sum_{n=1}^{N} h_n(x_n^*)$. In this case, a nonzero duality gap between g^* and f^* is obviously seen.

The significance of "time-sharing" condition results from the fact that it is always satisfied in multi-carrier/fading-based optimization problems in the limit sense when the number of carriers (fading states) N goes to infinity. Following the similar analysis given in [81], we now show that (P1) satisfies the "time-sharing" condition. Let the optimum value of (P1) given the APC P_{avg} and the average harvested power constraint $\bar{Q}$ be $f_1(P_{\text{avg}}, \bar{Q})$ and meanwhile let $\{p_x(\nu), \alpha_x(\nu)\}$ and $\{p_y(\nu), \alpha_y(\nu)\}$ denote the optimal solutions corresponding to $f_1(P_{\text{avg,x}}, \bar{Q}_x) = E_\nu[X_x(\nu)]$ and $f_1(P_{\text{avg,y}}, \bar{Q}_y) = E_\nu[X_y(\nu)]$, respectively. Considering to interleave the solutions $(p_x(\nu), \alpha_x(\nu))$ and $(p_y(\nu), \alpha_y(\nu))$ with a proportionality θ, denoted by $p_z(\nu) = \theta p_x(\nu) + (1-\theta) p_y(\nu)$, and $\alpha_z(\nu) = \theta \alpha_x(\nu) + (1-\theta)\alpha_y(\nu)$, within each fading state ν, we have the achievable objective function of (P1) given by

$$\begin{aligned} E_\nu[X_z(\nu)] =& E_\nu[\theta X_x(\nu) + (1-\theta) X_y(\nu)] \\ =& \theta E_\nu[X_x(\nu)] + (1-\theta) E_\nu[X_y(\nu)] \\ =& \theta f_1(P_{\text{avg,x}}, \bar{Q}_x) + (1-\theta) f_1(P_{\text{avg,y}}, \bar{Q}_y). \end{aligned} \tag{3.17}$$

On the other hand, it is easily to verify that

$$E_\nu[p_z(\nu)] = E_\nu[\theta p_x(\nu) + (1-\theta)p_y(\nu)] \leq \theta P_{\text{avg,x}} + (1-\theta)P_{\text{avg,y}}, \tag{3.18}$$

and similarly

$$E_\nu[Q(p_z(\nu))] = E_\nu[\theta Q(p_x(\nu)) + (1-\theta)Q(p_y(\nu))] \geq \theta\bar{Q}_x + (1-\theta)\bar{Q}_y. \tag{3.19}$$

Note that the approximation is introduced in (3.17) where $X(\nu)$ is assumed to remain constant within the considered fading block, which is however, automatically satisfied under the block fading assumption given in Section 3.3. Combining (3.17), (3.18), and (3.19), (P1) is thus proved to satisfy the "time-sharing" condition and therefore admits zero duality gap.

3.5.2 Optimal Solution to Secrecy Outage Probability Minimization

The Lagrangian of (P1) is expressed as

$$\begin{aligned}
&L(\{p(\nu)\},\{\alpha(\nu)\},\lambda,\mu) \\
&= E_\nu[X(\nu)] + \lambda(E_\nu[p(\nu)] - P_{\text{avg}}) - \mu(E_\nu[Q(p(\nu))] - \bar{Q}) \\
&= E_\nu[X(\nu) + \lambda p(\nu) - \zeta\mu g(\nu)p(\nu)] - \lambda P_{\text{avg}} + \mu\bar{Q},
\end{aligned} \tag{3.20}$$

where λ and μ are the dual variables associated with the APC, P_{avg}, and the average harvested power constraint, $\bar{Q}$, respectively. Then the (partial) Lagrange dual function of (P1) is expressed as

$$g(\lambda,\mu) = \min_{\{p(\nu)\leq P_{\text{peak}}\},\{\alpha(\nu)\in[0,1]\}} L(\{p(\nu)\},\{\alpha(\nu)\},\lambda,\mu). \tag{3.21}$$

The dual problem of (P1) is thus given by

$$(\text{P1}-\text{dual}): \ \underset{\lambda,\mu}{\texttt{Maximize}} \quad g(\lambda,\mu)$$
$$\texttt{Subject to} \quad \lambda \geq 0,\ \mu \geq 0.$$

The minimization problem in (3.21) can be decoupled into parallel subproblems each for one fading state all having the same structure. Specifically, for one particular fading state ν, define $L_1(p,\alpha) = X + \lambda p - \zeta\mu g p$. Then the associated subproblem given a pair of λ and μ is expressed as

$$(\text{P1}-\text{sub}): \ \underset{p,\alpha}{\texttt{Minimize}} \quad L_1(p,\alpha)$$
$$\texttt{Subject to} \quad p \leq P_{\text{peak}},$$
$$0 \leq \alpha \leq 1.$$

Note that we have dropped the index ν in $p(\nu)$, $\alpha(\nu)$ and $X(\nu)$ for brevity.

Given any $0 \leq \alpha \leq 1$, let $p_1(\alpha)$ denote the minimum required power to maintain a target secrecy rate r_0, i.e., $R(\alpha,p) \geq r_0$, it can be shown that

$$p_1(\alpha) = \begin{cases} \frac{-(\alpha\sigma_1^2 g + (1-\alpha)\sigma_2^2 h - 2^{r_0}\sigma_1^2\sigma_2^2 g) + \sqrt{\Delta}}{2\bar{\alpha}(1-\bar{\alpha})hg} & \text{if } 0 < \alpha < 1, \\ (2^{r_0}-1) \,/\, (\frac{h}{\sigma_1^2} - \frac{2^{r_0} g}{\sigma_2^2}) \text{ if } \alpha = 0 & \text{and } h > \frac{\sigma_1^2 2^{r_0} g}{\sigma_2^2}, \\ +\infty & \text{otherwise,} \end{cases} \tag{3.22}$$

where Δ is given by

$$\begin{aligned} \Delta &= \big(\alpha\sigma_1^2 g + \sigma_2^2(1-\alpha)h\big)^2 + 2^{r_0}\big(2^{r_0}\sigma_1^4 g^2 - 2\alpha\sigma_1^4 g^2 \\ &= +(-4\alpha^2 + 6\alpha - 2)\sigma_1^2\sigma_2^2 hg\big). \end{aligned} \tag{3.23}$$

Moreover, define $\tilde{\alpha}$ as the optimal solution to the following problem:

$$(\text{P1}-\text{search}): \ \underset{\alpha}{\texttt{Minimize}} \quad p_1(\alpha)$$
$$\texttt{Subject to} \quad 0 \le \alpha \le 1,$$

which can be obtained by a simple one-dimension search. Then we have the following proposition.

Proposition 3.5.1. *The optimal power allocations and power splitting ratios to problem (P1-sub) are given as*

$$\begin{cases} p^* = P_{\text{peak}},\ \alpha^* = \begin{cases} \tilde{\alpha} & \text{if } p_1(\tilde{\alpha}) \le P_{\text{peak}}, \\ 0 & \text{if } p_1(\tilde{\alpha}) > P_{\text{peak}}, \end{cases} & \text{if } g > \frac{\lambda}{\zeta\mu} \\ p^* = p_1(\tilde{\alpha}),\ \alpha^* = \tilde{\alpha}, \text{if } g \le \frac{\lambda}{\zeta\mu} \text{ and } p_1(\tilde{\alpha}) \le \min\left(\frac{1}{\lambda-\zeta\mu g}, P_{\text{peak}}\right), \\ p^* = 0 \qquad \alpha^* = 0, \text{ otherwise.} \end{cases} \tag{3.24}$$

Proof. Please refer to Appendix A. □

Remark 3.5.1. *We can draw some useful insight from Proposition 3.5.1 for the optimal power control policy for a given pair of* (λ, μ)*. When* $g > \frac{\lambda}{\zeta\mu}$*, which means a relatively better channel condition for the ER, the Tx needs to transmit with peak power in order to maximize the harvested energy at the ER. Under this circumstance, if furthermore,* $p_1(\tilde{\alpha}) > P_{\text{peak}}$*, i.e., the outage event is inevitable, there is no need to optimize* α *and thus it is set to be zero for simplicity; however, if* $p_1(\tilde{\alpha}) \le P_{\text{peak}}$*, the outage can be avoided by setting* α *to be any value satisfying* $p_1(\alpha) \le P_{\text{peak}}$*, and thus we set* $\alpha = \tilde{\alpha}$*. On the other hand, when* $g \le \frac{\lambda}{\zeta\mu}$*, we need to decide for the Tx whether to transmit with power* $p_1(\tilde{\alpha})$ *with power splitting ratio* $\tilde{\alpha}$*, or to shut down its transmission to save power, based on whether* $p_1(\tilde{\alpha})$ *is smaller or larger than a certain threshold, i.e.,* $\min(\frac{1}{\lambda-\zeta\mu g}, P_{\text{peak}})$*.*

According to Proposition 3.5.1, with a given pair of (λ, μ), (P1-sub) can be efficiently solved state by state based on (3.24). Problem (P1) is then iteratively

solved by updating (λ, μ) via the ellipsoid method [84], for which the details are omitted for brevity. Notice that the required sub-gradient for updating (λ, μ) can be shown to be $\big(E_\nu[p^*(\nu)] - P_{\text{avg}}, \bar{Q} - E_\nu[Q(p^*(\nu))]\big)$, where $p^*(\nu)$ is the optimal solution to problem (P1-sub) with given λ and μ.

3.5.3 Suboptimal Solution to Secrecy Outage Probability Minimization

Note that the optimal solution given in Proposition 3.5.1 requires an exhaustive search over α in (P1-search) for $\tilde{\alpha}$ in each of the fading states. In this subsection, we propose a suboptimal algorithm to solve (P1) with lower complexity based on the principle of alternating optimization. Specifically, by fixing $\alpha(\nu) = \bar{\alpha}(\nu)$, $\forall \nu$, we first optimize $\{p(\nu)\}$ by solving the following problem.

$$\begin{aligned} (\text{P1.1}): \ \underset{\{p(\nu)\}}{\texttt{Minimize}} \quad & E_\nu\left[X(\nu)\right] \\ \texttt{Subject to} \quad & E_\nu[p(\nu)] \le P_{\text{avg}}, \\ & p(\nu) \le P_{\text{peak}}, \ \forall \nu, \\ & E_\nu[Q(p(\nu))] \ge \bar{Q}. \end{aligned}$$

Let the optimal solution to (P1.1) be denoted by $\{\bar{p}(\nu)\}$, with $p(\nu) = \bar{p}(\nu)$, $\forall \nu$, we then optimize $\{\alpha(\nu)\}$ by solving the following problem.

$$\begin{aligned} (\text{P1.2}): \ \underset{\{\alpha(\nu)\}}{\texttt{Minimize}} \quad & E_\nu\left[X(\nu)\right] \\ \texttt{Subject to} \quad & 0 \le \alpha(\nu) \le 1, \ \forall \nu. \end{aligned}$$

The above procedure is repeated until both $\{p(\nu)\}$ and $\{\alpha(\nu)\}$ converge. In the following, we solve (P1.1) and (P1.2), respectively.

Problem (P1.1) is a non-convex problem since the objective function is not concave over $p(\nu)$. However, similar to (P1), it satisfies the "time-sharing" condition,

and thus we can use Lagrange duality method to solve it approximately with zero duality gap. Similarly as for problem (P1), problem (P1.1) can be decoupled into parallel subproblems each for one particular fading state and expressed as (by ignoring the fading state ν)

$$(\text{P1.1}-\text{sub}): \quad \underset{p}{\texttt{Minimize}} \quad L_1(p)$$
$$\texttt{Subject to} \quad p \le P_{\text{peak}},$$

where $L_1(p) = X + \lambda p - \zeta\mu g p$.

Through the similar analysis as for Proposition 3.5.1, given any $0 \le \bar{\alpha} \le 1$, the optimal solution to problem (P1.1-sub) is given as

$$p^* = \begin{cases} P_{\text{peak}} & \text{if } g > \frac{\lambda}{\zeta\mu}, \\ p_1(\bar{\alpha}) & \text{if } g \le \frac{\lambda}{\zeta\mu} \text{ and } p_1(\bar{\alpha}) \le \min\left(\frac{1}{\lambda-\zeta\mu g}, P_{\text{peak}}\right), \\ 0 & \text{otherwise.} \end{cases} \tag{3.25}$$

With a given pair of (λ, μ), (P1.1-sub) can be efficiently solved state by state based on (3.25). Problem (P1.1) can thus be iteratively solved by updating (λ, μ) via the ellipsoid method.

Next, we derive the optimal power splitting ratios $\{\alpha(\nu)\}$ for problem (P1.2) with given $\{\bar{p}(\nu)\}$. Note that the objective function of (P1.2) is separable over different fading states of ν. Hence, we only need to solve the following problem for each of the fading states.

$$\begin{aligned} &\underset{\alpha}{\texttt{Minimize}} \quad X \\ &\texttt{Subject to} \quad 0 \le \alpha \le 1. \end{aligned} \tag{3.26}$$

Note that we have dropped the index ν for brevity.

Define $\Phi = \{\alpha | R(\alpha, \bar{p}) \ge r_0\}$ as the set of α that can guarantee the non-outage secrecy information transmission given $\bar{p}$. If $\Phi = \emptyset$, the outage cannot be avoided and

thus any $0 \leq \alpha \leq 1$ can be the optimal solution to problem (3.26). Otherwise, any $\alpha \in \Phi$ is optimal to problem (3.26). To select the best solution among the feasible α's, we solve the following problem.

$$(\text{P1.2} - \text{sub}): \quad \underset{\alpha}{\texttt{Maximize}} \quad R(\alpha, \bar{p})$$
$$\texttt{Subject to} \quad 0 \leq \alpha \leq 1.$$

Define $x = \frac{\sigma_1^2}{h\bar{p}} - \frac{\sigma_2^2}{g\bar{p}}$. Then we have the following proposition.

Proposition 3.5.2. *If Φ is non-empty, the optimal solution to problem (P1.2-sub) is given by*

$$\hat{\alpha}^* = \begin{cases} 0 & x < -1, \\ \frac{1}{2} + \frac{x}{2} & -1 \leq x < 1, \\ 1 & x \geq 1. \end{cases} \tag{3.27}$$

Proof. Please refer to Appendix B. □

By combining both the cases of $\Phi \neq \emptyset$ and $\Phi = \emptyset$, the optimal solution to problem (P1.2-sub) is given by $\alpha^* = \hat{\alpha}^*$. Hence, problem (P1.2) for all ν's can be solved according to (3.27).

With both problems (P1.1) and (P1.2) solved, we can then iteratively solve the two problems to obtain a suboptimal solution for (P1). It is worth noting that the suboptimal algorithm proposed guarantees that the outage probability obtained is non-increasing after each iteration; thus the algorithm is ensured to at least converge to a locally optimal solution to (P1).

3.6 Proposed Solutions for No-Delay-Limited Case

In this section, we propose both optimal and suboptimal solutions to solve (P2).

3.6.1 Optimal Solution to ESC Maximization

First, we propose an optimal algorithm to solve (P2). Similar to Section 3.5.2, based on the Lagrange duality method, problem (P2) can be decoupled into parallel subproblems all having the same structure and each for one fading state. Specifically, for one particular fading state ν, we define $L_2(p,\alpha) = R(\alpha,p) - \lambda p + \zeta\mu g p$, where $R(\alpha,p)$ is given in (3.10). Then the associated subproblem to solve for fading state ν is expressed as

$$\begin{aligned}(\text{P2}-\text{sub}):\ \underset{p,\alpha}{\texttt{Maximize}} \quad & L_2(p,\alpha)\\ \texttt{Subject to} \quad & p \le P_{\text{peak}},\\ & 0 \le \alpha < 1.\end{aligned}$$

Note that we have dropped the index ν in $p(\nu)$ and $\alpha(\nu)$ for brevity.

Since $R(\alpha,p)$ is not concave over p and α, problem (P2-sub) is non-convex and thus difficult to be solved by applying convex optimization techniques. Hence, we propose a two-stage procedure to solve (P2-sub) optimally. First, we fix $\alpha = \bar{\alpha}$ and then solve (P2-sub) to find the corresponding optimal power allocation $\bar{p}$. Let $f_\nu(\bar{\alpha})$ denote the optimal value of (P2-sub) given $\alpha = \bar{\alpha}$. Next, the optimal α^* to (P2-sub) is obtained by $\max\limits_{0\le\bar{\alpha}\le 1} f_\nu(\bar{\alpha})$, which can be solved by a one-dimension search over $\bar{\alpha} \in [0,1]$. Therefore, in the following we focus on how to solve problem (P2-sub) with $\alpha = \bar{\alpha}$. First, we obtain the derivative of $L_2(p,\bar{\alpha})$ over p as

$$\frac{\partial L_2(p,\bar{\alpha})}{\partial p} = \begin{cases} \frac{Ap^3+Bp^2+Cp+D}{E} & \text{if } p > \frac{\sigma_1^2}{\bar{\alpha}h} - \frac{\sigma_2^2}{\bar{\alpha}g},\\ -\lambda + \mu\zeta g & \text{otherwise,}\end{cases} \tag{3.28}$$

where $A \triangleq \bar{\alpha}hg^2(\lambda-\mu\zeta g)(\bar{\alpha}-1)\ln 2$, $B \triangleq h(\bar{\alpha}-1)F - \bar{\alpha}hg^2(\bar{\alpha}-1) - \bar{\alpha}g^2\sigma_1^2(\lambda-\mu\zeta g)\ln 2$, $C \triangleq h\sigma_2^4(\lambda-\mu\zeta g)(\bar{\alpha}-1)\ln 2 - \sigma_1^2 F - hg\sigma_2^2(\bar{\alpha}-1)^2 - (hg\sigma_2^2 + \bar{\alpha}hg\sigma_2^2)(\bar{\alpha}-1)$, $D \triangleq g\sigma_2^2\sigma_1^2(\bar{\alpha}-1) - h\sigma_2^4(\bar{\alpha}-1) - \sigma_2^4\sigma_1^2(\lambda-\mu\zeta g)\ln 2$, $E \triangleq (\sigma_1^2 + (1-\bar{\alpha})ph)(\sigma_2^2+\bar{\alpha}pg)(\sigma_2^2 + pg)\ln 2$, and $F \triangleq g\sigma_2^2(\lambda-\mu\zeta g)(1+\bar{\alpha})\ln 2$. It can be observed from (3.28) that the

monotonicity of $L_2(p, \bar{\alpha})$ closely relates to the following cubic equation:

$$Ap^3 + Bp^2 + Cp + D = 0. \tag{3.29}$$

According to fundamental theorem of algebra, there are at most three roots (counted with multiplicity) to (3.29), denoted by $x_1, x_2,$ and, x_3. Define a set as $\mathcal{X} = \{x_1, x_2, x_3\}$. Since $p \in \mathbb{R}$, only real roots in $\mathcal{X}$ need to be taken into account. Thus, we define another set Ψ as follows:

$$\Psi = \{x | x \in \mathbb{R}, 0 \leq x \leq P_{\text{peak}}, x \in \mathcal{X}\} \cup \{0, P_{\text{peak}}\}, \tag{3.30}$$

where $2 \leq |\Psi| \leq 5$, with $|\cdot|$ denoting the cardinality of a set. Note that $|\Psi| = 2$ when no real roots fall in the interval $[0, P_{\text{peak}}]$, while $|\Psi| = 5$ when there are three distinct real roots in $(0, P_{\text{peak}})$. Next, it is easy to show that the optimal p that maximizes $L_2(p, \bar{\alpha})$ over $p \in [0, P_{\text{peak}}]$ is obtained via a simple search over Ψ, i.e.,

$$\bar{p}(\lambda, \mu) = \arg \max_{p \in \Psi} L_2(p, \bar{\alpha}). \tag{3.31}$$

As a result, problem (P2-sub) is solved given any pair of (λ, μ). Problem (P2) is then solved by iteratively updating (λ, μ) by the ellipsoid method.

3.6.2 Suboptimal Solution to ESC Maximization

Note that the optimal solution to (P2) requires a one-dimension search to find α^* for each fading state. Thus, in this subsection, we propose a suboptimal algorithm to solve (P2) with lower complexity based on alternating optimization. Specifically,

by fixing $\alpha(\nu) = \bar{\alpha}(\nu)$, $\forall \nu$, we first optimize $\{p(\nu)\}$ by solving the following problem.

$$\begin{aligned} (\text{P2.1}): \ \underset{\{p(\nu)\}}{\texttt{Maximize}} \quad & E_\nu \left[R(\bar{\alpha}(\nu), p(\nu)) \right] \\ \texttt{Subject to} \quad & E_\nu[p(\nu)] \leq P_{\text{avg}}, \\ & p(\nu) \leq P_{\text{peak}}, \ \forall \nu, \\ & E_\nu[Q(p(\nu))] \geq \bar{Q}. \end{aligned}$$

Let the optimal solution of (P2.1) be denoted by $\{\bar{p}(\nu)\}$. With $p(\nu) = \bar{p}(\nu)$, $\forall \nu$, we then optimize $\{\alpha(\nu)\}$ by solving the following problem.

$$\begin{aligned} (\text{P2.2}): \ \underset{\{\alpha(\nu)\}}{\texttt{Maximize}} \quad & E_\nu \left[R(\alpha(\nu), \bar{p}(\nu)) \right] \\ \texttt{Subject to} \quad & 0 \leq \alpha(\nu) \leq 1, \ \forall \nu. \end{aligned}$$

The above two-stage procedure is repeated until both $\{\bar{p}(\nu)\}$ and $\{\bar{\alpha}(\nu)\}$ converge. In the following, we solve (P2.1) and (P2.2), respectively.

Similar to (P1.1), problem (P2.1) can be decoupled into parallel subproblems each for one fading state and expressed as (by ignoring the fading state ν)

$$\begin{aligned} (\text{P2.1} - \text{sub}): \ \underset{p}{\texttt{Maximize}} \quad & L_2(p) \\ \texttt{Subject to} \quad & p \leq P_{\text{peak}}, \end{aligned}$$

where $L_2(p) = R(\bar{\alpha}, p) - \lambda p + \zeta \mu g p$.

Note that problem (P2.1-sub) is equivalent to problem (P2-sub) with given $\alpha = \bar{\alpha}$, the solution of which has been given in (3.31). As a result, problem (P2.1-sub) can be efficiently solved. Then, problem (P2.1) can be solved by iteratively updating (λ, μ) via the ellipsoid method.

Next, we derive the optimal power splitting ratios $\{\alpha(\nu)\}$ for problem (P2.2) with given $\{\bar{p}(\nu)\}$ obtained by solving problem (P2.1). Note that the objective function of (P2.2) is separable over different fading states. Thus, for each fading state ν, we

need to solve the following problem (by dropping the index ν for brevity):

$$(\text{P2.2}-\text{sub}): \underset{\alpha}{\texttt{Maximize}} \quad R(\alpha, \bar{p})$$
$$\texttt{Subject to} \quad 0 \leq \alpha \leq 1.$$

Note that problem (P2.2-sub) is the same as problem (P2.1-sub) in Section 3.5.3, the solution of which has already been derived in Proposition 3.5.2. Hence, problem (P2.2) for all ν's can be solved according to (3.27).

With both problems (P2.1) and (P2.2) solved, we can obtain a suboptimal solution for (P2) by iteratively solving these two problems. Similar to that for (P1), this suboptimal algorithm guarantees that the ESC is non-decreasing after each iteration, and thus convergence to at least a local optimal solution of (P2) is ensured.

3.7 Benchmark Schemes

In this section, we introduce two benchmark schemes, where no AN is used at the transmitter, and the AN is used but is unknown to both the IR and ER, respectively.

First, consider the case when no AN is employed, i.e., $\alpha(\nu) = 0, \forall \nu$ for both the delay-limited secrecy transmission and the non-delay-limited counterpart. In this case, the SNRs at the IR and ER at fading state ν given in (3.8) and (3.9) reduce to

$$\text{SNR}'_{\text{IR}}(\alpha(\nu), p(\nu)) = \frac{h(\nu)p(\nu)}{\sigma_1^2}, \tag{3.32}$$
$$\text{SNR}'_{\text{ER}}(\alpha(\nu), p(\nu)) = \frac{g(\nu)p(\nu)}{\sigma_2^2}, \tag{3.33}$$

respectively. Thus, the secrecy rate given in (3.10) reduces to

$$R'(p(\nu)) = \left[\log_2\left(1 + \frac{h(\nu)p(\nu)}{\sigma_1^2}\right) - \log_2\left(1 + \frac{g(\nu)p(\nu)}{\sigma_2^2}\right)\right]^+. \tag{3.34}$$

It follows from (3.34) that the outage probability becomes $\delta' = Pr(R'(p(\nu)) < r_0)$,

or equivalently, $\delta' = E_\nu[X'(\nu)]$, where $X'(\nu)$ is modified from (3.14) as

$$X'(\nu) = \begin{cases} 1 & \text{if } R'(p(\nu)) < r_0, \\ 0 & \text{otherwise.} \end{cases} \tag{3.35}$$

Thus, (P1) reduces to the following problem.

$$\begin{aligned} (\text{P1} - \text{NoAN}): \quad & \underset{\{p(\nu)\}}{\texttt{Minimize}} \quad E_\nu[X'(\nu)] \\ & \texttt{Subject to} \quad E_\nu[p(\nu)] \le P_{\text{avg}}, \\ & \qquad\qquad\quad\ p(\nu) \le P_{\text{peak}}, \forall \nu, \\ & \qquad\qquad\quad\ E_\nu[Q(p(\nu))] \ge \bar{Q}. \end{aligned}$$

Accordingly, (P2) reduces to the following problem.

$$\begin{aligned} (\text{P2} - \text{NoAN}): \quad & \underset{\{p(\nu)\}}{\texttt{Maximize}} \quad E_\nu[R'(p(\nu))] \\ & \texttt{Subject to} \quad E_\nu[p(\nu)] \le P_{\text{avg}}, \\ & \qquad\qquad\quad\ p(\nu) \le P_{\text{peak}}, \forall \nu, \\ & \qquad\qquad\quad\ E_\nu[Q(p(\nu))] \ge \bar{Q}. \end{aligned}$$

Note that (P1-NoAN) and (P2-NoAN) can be solved by simply setting $\alpha(\nu) = 0$ in (P1.1) and (P2.1), respectively.

Next, consider the case when the AN is used but is unknown to both the IR and ER, i.e., it cannot be canceled by the IR any more unlike that assumed in Sections 3.5 and 3.6. In this case, the SNR expression at the ER at fading state ν is unchanged as (3.9), while the SNR at the IR at fading state ν needs to be modified as

$$\text{SNR}''_{\text{IR}}(\alpha(\nu), p(\nu)) = \frac{(1-\alpha(\nu))h(\nu)p(\nu)}{\alpha(\nu)h(\nu)p(\nu) + \sigma_1^2}. \tag{3.36}$$

Then, the achievable secrecy rate given in (3.10) is modified accordingly as

$$R''(\alpha(\nu),p(\nu)) = \Bigg[\log_2\left(1+\frac{(1-\alpha(\nu))h(\nu)p(\nu)}{\alpha(\nu)h(\nu)p(\nu)+\sigma_1^2}\right) - \log_2\left(1+\frac{(1-\alpha(\nu))g(\nu)p(\nu)}{\alpha(\nu)g(\nu)p(\nu)+\sigma_2^2}\right)\Bigg]^+. \tag{3.37}$$

It follows from (3.37) that the outage probability reduces to $\delta'' = Pr(R''(\alpha(\nu),p(\nu)) < r_0)$, or equivalently, $\delta'' = E_\nu[X''(\nu)]$, where $X''(\nu)$ is also modified from (3.14) as

$$X''(\nu) = \begin{cases} 1 & \text{if } R''(\alpha(\nu),p(\nu)) < r_0, \\ 0 & \text{otherwise.} \end{cases} \tag{3.38}$$

Thus, (P1) is reformulated as

$$\begin{aligned} (\text{P1}-\text{NoCancel}): \quad & \underset{\{p(\nu),\alpha(\nu)\}}{\texttt{Minimize}} && E_\nu[X''(\nu)] \\ & \texttt{Subject to} && E_\nu[p(\nu)] \le P_{\text{avg}}, \\ & && p(\nu) \le P_{\text{peak}}, \forall\nu, \\ & && E_\nu[Q(p(\nu))] \ge \bar{Q}, \\ & && 0 \le \alpha(\nu) \le 1, \forall\nu. \end{aligned}$$

Accordingly, (P2) is reformulated as

$$\begin{aligned} (\text{P2}-\text{NoCancel}): \quad & \underset{\{p(\nu),\alpha(\nu)\}}{\texttt{Maximize}} && E_\nu[R''(\alpha(\nu),p(\nu))] \\ & \texttt{Subject to} && E_\nu[p(\nu)] \le P_{\text{avg}}, \\ & && p(\nu) \le P_{\text{peak}}, \forall\nu, \\ & && E_\nu[Q(p(\nu))] \ge \bar{Q}, \\ & && 0 \le \alpha(\nu) \le 1, \forall\nu. \end{aligned}$$

(P1-NoCancel) and (P2-NoCancel) are both non-convex problems because $X''(\nu)$ and $R''(\alpha(\nu),p(\nu))$ are non-convex and non-concave over $p(\nu)$ and $\alpha(\nu)$, respectively.

However, we have the following proposition on their optimal solutions.

Proposition 3.7.1. *The optimal solution to problem (P1-NoCancel) and (P2-NoCancel) must satisfy* $\alpha^*(\nu) = 0, \forall \nu$.

Proof. For problems (P1-NoCancel) and (P2-NoCancel), suppose that the average harvested power constraint is not present, the optimal power splitting ratios for both problems can be shown to be $\alpha^*(\nu) = 0, \forall \nu$, by solving $\max\limits_{0 \le \alpha(\nu) \le 1} R''(\alpha(\nu), \bar{p}(\nu))$ at each fading state ν (c.f. (3.37)), according to [85]. The reason is as follows. Since $\frac{\partial R''(\alpha,\bar{p})}{\partial \alpha} = \frac{-1}{\ln 2} \frac{(h\sigma_2^2 - g\sigma_1^2)\bar{p}}{(\alpha h\bar{p}+\sigma_1^2)(\alpha g\bar{p}+\sigma_2^2)} \le 0$, $R''(\alpha, \bar{p})$ is monotonically non-increasing w.r.t. α over the interval $[0, 1]$, and thus attains its maximum at $\alpha = 0$. Now, with the average harvested power constraint added, since the harvested power given in (3.11) in each fading state ν is independent of $\alpha(\nu)$, it is also true that setting $\alpha^*(\nu) = 0, \forall \nu$, has no loss of optimality. Combining the above two results, we conclude that $\alpha^*(\nu) = 0, \forall \nu$, should be optimal for both problems. Proposition 3.7.1 is thus proved. □

Proposition 3.7.1 indicates that no AN should be used in (P1-NoCancel) or (P2-NoCancel), if it cannot be canceled by the IR. As a result, (P1-NoCancel) and (P2-NoCancel) are equivalent to the previous two problems, (P1-NoAN) and (P2-NoAN), respectively, which can be efficiently solved.

3.8 Numerical Results

In this section, we provide numerical examples to evaluate the performance of our proposed optimal and suboptimal algorithms in Sections 3.5 and 3.6, against the two benchmark schemes introduced in Section 3.7. For comparison, we also consider the following heuristic approach to solve (P1) and (P2). First, we fix $\alpha(\nu) = \bar{\alpha}$, $\forall \nu$, in (P1) or (P2), i.e., a uniform power splitting ratio for all fading states is assumed; then, we solve (P1.1) or (P2.1) to obtain the optimal $\{p(\nu)\}$. For convenience, in the sequel we refer to the above scheme as Fixed-$\bar{\alpha}$. Compared with the two suboptimal algorithms proposed in Sections 3.5 and 3.6, which require iteratively updating between $\{\alpha(\nu)\}$ and $\{p(\nu)\}$ until their convergence, the algorithm of Fixed-$\bar{\alpha}$

with fixed $\alpha(\nu) = \bar{\alpha}, \forall \nu$, only needs one-shot for solving $\{p(\nu)\}$, and thus has a much lower complexity.

We set $P_{\text{avg}} = 100\text{mW}$ or 20dBm, $P_{\text{peak}} = 1\text{W}$ or 30dBm, $\zeta = 50\%$, and $\sigma_1^2 = \sigma_2^2 = -50\text{dBm}$. The distance-dependent pass loss model is given by

$$L = A_0 \left(\frac{d}{d_0}\right)^{-\alpha}, d \geq d_0, \tag{3.39}$$

where A_0 is set to be 10^{-3}, d denotes the distance between the Tx to the IR or ER, d_0 is a reference distance set to be 1m, and α is the path loss exponent set to be 3. It is assumed that $h(\nu)$ and $g(\nu)$ are independent exponentially distributed RVs (accounting for short-term Rayleigh fading) with their average power values specified by (3.39).

3.8.1 Secrecy Outage-Energy Trade-off

At first, we consider (P1) for characterizing the trade-offs between the secrecy outage probability for the IR and the average harvested power for the ER. Specifically, we adopt the (secrecy) O-E region [81], which consists of all the pairs of achievable (secrecy) non-outage probability ϵ and average harvested power E for a given set of P_{avg} and P_{peak}, which is defined as

$$\mathcal{C}_{\text{O-E}} \triangleq \bigcup_{\substack{E_\nu[p(\nu)] \leq P_{\text{avg}} \\ p(\nu) \leq P_{\text{peak}}, \forall \nu \\ 0 \leq \alpha(\nu) \leq 1, \forall \nu}} \Big\{(\epsilon, E) : \epsilon \leq 1 - \delta, E \leq Q_{\text{avg}}\Big\}, \tag{3.40}$$

where Q_{avg} is given in (3.12), and $1 - \delta$ is the non-outage probability w.r.t. a given secrecy rate r_0, where δ is given in (3.13). Note that by solving (P1) with different $\bar{Q}$'s, the boundary of the corresponding O-E region for each considered scheme can be obtained accordingly.

Consider a setup where the IR and the ER are of an identical distance of 2m to the Tx. The target secret rate is set as $r_0 = 6.5\text{bps/Hz}$. Fig. 3.3 shows the O-E regions of the different schemes. It is observed that compared with both the

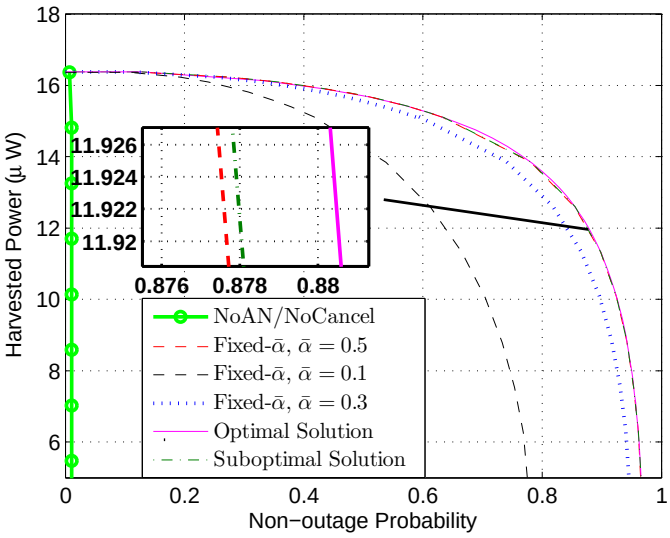


Figure 3.3: Achievable O-E regions with a target secret rate $r_0 = 6.5$bits/sec/Hz by different power allocation schemes when the IR and ER are both 2m away from the Tx.

schemes of NoAN and NoCancel, the proposed optimal algorithm with the use of AN achieves substantially improved O-E trade-offs thanks to the AN cancellation at the IR. For example, when an average harvested power of 7.0μW is achieved, the secrecy outage probability can be made less than 5% versus more than 98%. Furthermore, it is observed that when the AN can be canceled by the IR, the O-E region achieved by the suboptimal solution with alternating optimization is very close to that of the optimal solution. Furthermore, it is also observed that the O-E region achieved by Fixed-$\bar{\alpha}$ with $\bar{\alpha} = 0.5, \forall \nu$, has only negligible loss as compared to that of the optimal solution. The reason is as follows. In this setup, both the IR and the ER are very close to the Tx, and thus their average SNRs are high. It thus follows from (3.27) that when SNRs for the IR and the ER are high enough, $x = \frac{\sigma_1^2}{h\bar{p}} - \frac{\sigma_2^2}{g\bar{p}}$ tends to be zero, and as a result, if the transmission is on, i.e., $\bar{p} \neq 0$, the optimal power splitting ratios to (P1.2) becomes $\alpha^*(\nu) \approx 0.5$, $\forall \nu$. Last, we observe that the O-E trade-offs achieved by Fixed-$\bar{\alpha}$ with other fixed values of $\bar{\alpha}$ instead of $\bar{\alpha} = 0.5$ deviate more notably from that of the optimal solution.

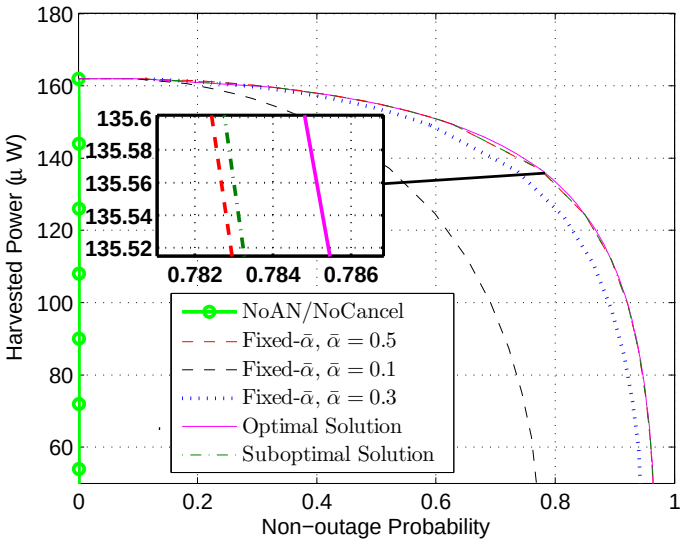


Figure 3.4: Achievable O-E regions with a target secret rate $r_0 = 6.5$bits/sec/Hz by different power allocation schemes when the IR and ER are 2m and 1m away from the Tx, respectively.

Next, we consider a more challenging setup for secrecy transmission when the ER is in more proximity to the Tx than the IR. Specifically, we assume that the IR and ER are 2m and 1m away from the Tx, respectively. Fig. 3.4 shows the O-E regions achieved by different schemes. Compared with Fig. 3.3, it is observed that despite of the much worse channel condition for the IR than the ER, the achieved outage probability for secrecy transmission is almost unchanged. Also note from Fig. 3.4 that the achievable average harvested power for the ER is as about 10 times as that in Fig. 3.3. However, it is observed that under this setup, the outage probability achieved by the schemes of NoAN or NoCancel is almost one due to the severely deteriorated average SNR of the IR's channel.

3.8.2 Secrecy Rate-Energy Trade-off

Next, we consider (P2) for characterizing the trade-offs between the ESC for the IR and the average harvested power for the ER. Specifically, we adopt the (secrecy) R-E region [82], which consists of all the pairs of achievable (secrecy) rate R and

harvested power E for a given set of P_{avg} and P_{peak}, which is defined as

$$\mathcal{C}_{\text{R-E}} \triangleq \bigcup_{\substack{E_\nu[p(\nu)]\le P_{\text{avg}} \\ p(\nu)\le P_{\text{peak}} \\ 0\le\alpha(\nu)\le 1}} \Big\{(R,E): R \le C_s, E \le Q_{\text{avg}}\Big\}, \tag{3.41}$$

where Q_{avg} is given in (3.12), and C_s is expressed as $C_s = E_\nu[R(\nu)]$, with $R(\nu)$ given in (3.10), (3.34) and (3.37), respectively, for different schemes. Note that by solving (P2) with different $\bar{Q}$'s, the boundary of the corresponding R-E region for each considered scheme can be obtained.

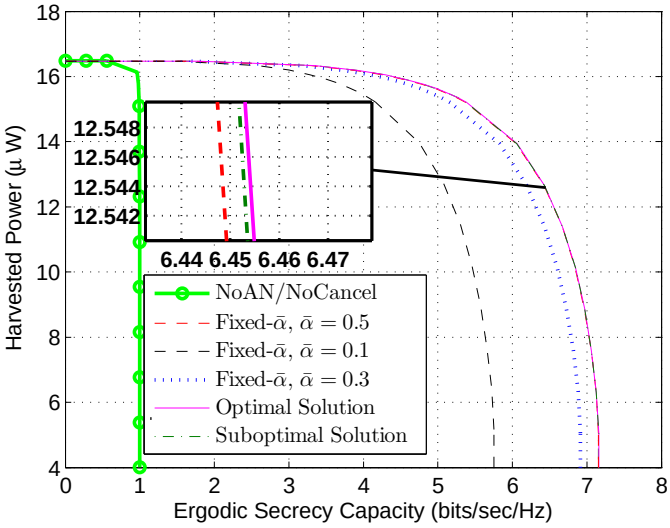


Figure 3.5: Achievable R-E regions by different power allocation schemes when the IR and ER are both 2m away from the Tx.

Similar to the case of O-E region, we first consider the setup when the IR and the ER are of an identical distance of 2m to the Tx. Fig. 3.5 shows the R-E regions of the different schemes. It is observed that compared with the scheme of NoAN (or NoCancel), the proposed AN-aided optimal solution achieves substantially improved R-E trade-offs due to the cancelable AN at the IR. For example, when an average harvested power of 6μW is achieved, the ESC is increased by about 700%. Furthermore, it is observed that when the AN can be canceled by the IR,

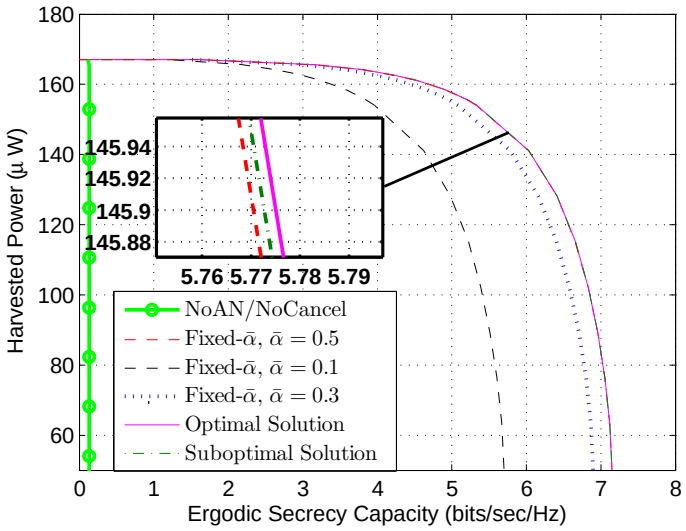


Figure 3.6: Achievable R-E regions by different power allocation schemes when the IR and ER are 2m and 1m away from the Tx, respectively.

the R-E region achieved by the suboptimal solution is very close to that by the optimal solution. Finally, similar to the case of O-E region, the R-E region achieved by Fixed-$\bar{\alpha}$ with $\bar{\alpha} = 0.5, \forall \nu$, is the best compared with those achieved by other fixed values of $\bar{\alpha}$, i.e., $\bar{\alpha} = 0.1$ and $\bar{\alpha} = 0.3$.

Next, we consider the same setup with unequal distances from the Tx to the ER and IR as for Fig. 3.4. Fig. 3.6 shows the R-E regions achieved by different schemes. Compared to Fig. 3.5, it is observed that the performance gaps between the proposed optimal/suboptimal solutions and the scheme of NoAN or NoCancel become more substantial.

3.9 Chapter Summary

The important issue of PLS in emerging SWIPT applications was studied in this chapter. Under a simplified three-node SISO fading wiretap channel setup, a dual use of the AN was proposed for both interfering with and transferring energy to the ER, under the assumption that the AN is perfectly canceled at the IR. The transmit power

allocations and power splitting ratios over the fading channel were jointly optimized to minimize the outage probability for delay-limited secrecy transmission, and to maximize the average rate for no-delay-limited secrecy transmission, respectively, subject to the combined average and peak power constraint at the Tx, as well as an average EH constraint at the ER. Optimal solutions to these non-convex problems were derived, and suboptimal solutions of lower complexity were also proposed based on the alternating optimization. Through extensive simulation results, the proposed schemes were shown to achieve considerable (secrecy) Outage-Energy (O-E) and (secrecy) Rate-Energy (R-E) trade-off gains, as compared to the schemes without the use of AN.

Chapter 4

HJ-aided AF Relaying for Secrecy in SWIPT Networks

4.1 Introduction

It was assumed in Chapter 3 that the ERs in the SWIPT systems attempt to intercept the information for the IR, which might be overly protective at times, since it is possible that some ERs are, however, helpful from the perspective that they cooperatively contribute to communications utilizing their wirelessly harvested energy. It is well known that in the classical three-node SISO wiretap channel, non-zero secrecy capacity is achievable if and only if the eavesdropper sees a degraded channel of the main one seen by the legitimate Rx. This bottleneck has been finally broken through from the signal-processing perspective by exploiting CJ, which selectively interferes with the legitimate Rx and the eavesdropper, respectively, such that the equivalent eavesdropper's channel becomes a degraded version of the main one. The effect of CJ is also deemed further enhanced by taking advantage the multi-antenna array gain. In spite of being a promising solution to break the "degraded channel" assumption, the technique of CJ is nevertheless prohibitive due to the extra supply of power for external helpers, especially when they are ultra-low-power applications with limited battery capacity. Hence, following the recent advances in WPCN [71], a self-sustaining *harvest-and-jam (HJ)* relaying protocol is proposed in this chapter, where in the first transmission phase a single-antenna Tx transfers confidential information to a multiple-antenna AF relay and simultaneously power to a group of multi-antenna EH-enabled ERs, while in the

second phase, the relay amplifies and forwards the information to the IR under the protection of the AN generated by the ERs merely using their harvested energy from the received signals in the first transmission phase. In particular, we study the use of multi-antenna HJ helpers in a multi-antenna amplify-and-forward (AF) relay wiretap channel assuming that the direct link between the source and destination is broken. The goal is to maximize the achievable secrecy rate at the destination subject to the transmit power constraints of the AF relay and HJ helpers. In the case of perfect CSI, the joint optimization of the artificial noise (AN) covariance for jamming and the AF beamforming matrix as well as suboptimal solutions with lower complexity are presented all based on SDR, which proves to be tight in this case. Under practical circumstances where the CSI is imperfect, the formulation of the robust optimization is provided for maximizing the worst-case secrecy rate. Using SDR techniques, a near-optimal robust scheme is also proposed. Numerical results are given to validate the effectiveness of the HJ protocol.

The rest of the chapter is organized as follows. The AF relaying SWIPT system model is introduced and the HJ protocol is described in Section 4.3. In Section 4.4, the secrecy rate maximization problem with perfect CSI is presented and the joint-optimal AN covariance and AF beamforming matrices as well as a complexity-reduced suboptimal solution is given. Effective solutions are proposed in Section 4.5 to tackle the case of imperfect CSI via the worst-case robust formulation. In Section 4.6, numerical results are provided to compare different schemes. Finally, the chapter is concluded in Section 4.7.

4.2 Related Work

4.2.1 Cooperation Strategies for PLS

PLS issues in the rapidly growing cooperative networks have attracted much attention. Cooperative approaches, such as CJ communications, have been widely examined [7, 8, 56, 86]. The idea is to assist the transmitter in the secrecy transmission

by generating an AN to interfere with the eavesdropper via either multiple antennas or external trusted helpers [21, 36, 87, 88]. However, all of those utilizing ANs require additional supply of power and therefore incur extra system costs. Meanwhile, collaborative use of relays to form effective beams jamming the eavesdropper, i.e., *secure collaborative relay beamforming*, has been studied for relay-wiretap channels with single eavesdropper in [45], multiple eavesdroppers with AF relays and DF relays in [46] and [35], respectively. All, however, assumed the availability of perfect channel state information (CSI).

4.2.2 (Worst-Case) Robust Secrecy Optimization

The assumption of perfect CSI of the eavesdroppers appears to be too ideal because the eavesdroppers, despite likely being subscribed users, wish to hide from the transmitter without being cooperative in the stage of channel estimation. Even if they are bound to help the transmitter in obtaining their CSIs to facilitate their own communication, the CSIs at the transmitter side will change due to mobility and Doppler effect, and may be outdated. Moreover, even for the legitimate users, the estimated CSIs may also be subject to quantization errors due to the limited capacity of the feedback channel, although the inaccuracy is reasonably assumed less severe than that for the eavesdroppers. To tackle this issue, state-of-art schemes have been developed ([89] and the references therein), among which the *worst-case secrecy rate* is commonly employed to formulate the robust secrecy rate maximization problem [29, 41, 49, 58, 88]. The robust transmit covariance design for the secrecy rate maximization in a MISO channel overheard by multi-antenna eavesdroppers was considered in [49] while the enhanced secrecy performance was achieved by introducing a friendly jammer in the same scenario in [29], in which a joint optimization of the robust transmit covariance and power allocation between the source and the helper was studied via geometric programming. More recently, [58] studied a joint robust design of the information beams, the AN and the energy signals for SWIPT networks with quality-of-service (QoS) constraints.

Note that [48] proposed robust AF relay beamforming against the eavesdropper's channel though, the solutions were yet suboptimal. Furthermore, of particular relevance to this chapter is [51] that jointly optimized the AF beamforming matrices and AN covariances in a relay wiretap channel with multiple multi-antenna AF relays and multiple multi-antenna eavesdroppers via a worst-case robust formulation. While their network model is similar to the concerned one in the following, the difference between this chapter and [51] is twofold. On one hand, the AN generated by the friendly jammers in this chapter are subject to their respective channels from the transmitter during WPT in the first transmission phase. On the other hand, the technique in [51, *Proposition 1*] cannot be applied herein since the AN beams and the forwarded information are now transmitted via separate channels. As a consequence, to the best of authors' knowledge, the proposed worst-case based *robust* optimization scheme that incorporates imperfect CSIs into all the HJ helpers, has not yet been addressed in the literature.

4.2.3 Wireless Powered CJ

It is worth noting that devising a wireless-powered friendly jammer to enhance PHY-layer security for a direct transmission protocol was studied in [72], in which the "harvesting" blocks and "jamming" blocks were well exploited to compose four different types of harvesting-jamming cycles. Compared to [72], which focused on the dedicated scheduling of "harvest" and "jam" operations and its long-term performance, this chapter is concerned with adaptive rate/power optimization with multiple HJ helpers to achieve higher (worst-case) secrecy rate. Moreover, instead of assuming perfect channels to/from the HJ helpers, the *robust* optimization algorithms proposed in this chapter take imperfect CSI of legitimate channels into account as well.

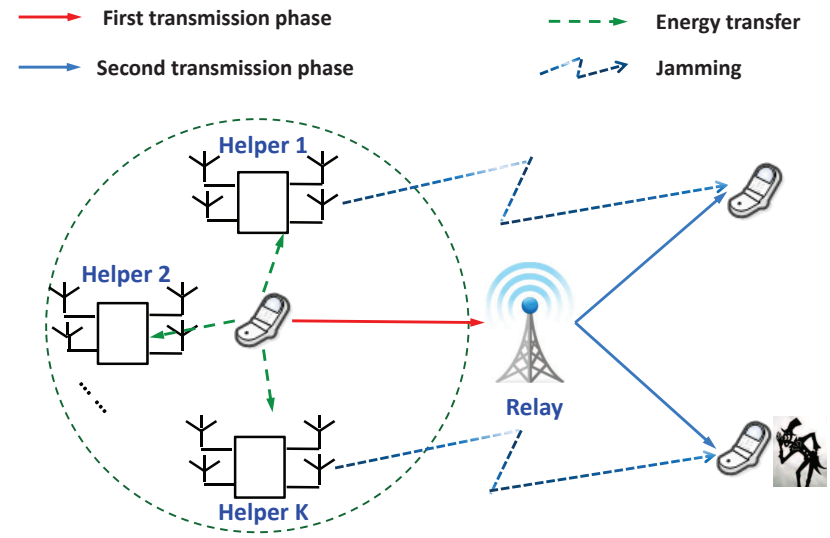
4.3 System Model

We consider a cooperative relay wiretap channel for SWIPT over a given frequency band as shown in Fig. 4.1(a). We assume that there is a transmitter, named Alice, sending confidential messages to the IR, Bob, in the presence of an eavesdropper [90], Eve, with the aid of a multi-antenna AF relay and K ERs willing to act as HJ helpers, $\mathcal{H}_{\text{helper}} = \{\mathsf{H}_1, \ldots, \mathsf{H}_K\}$. The transmitter, ERs, and the AF relay are deployed in a same cluster that is relatively far away from the destination and Eve, such that there is no direct link from the transmitter to the receiver or Eve, respectively [45, 46]. Moreover, the ERs are assumed to be located closer to the transmitter than the AF relay in order that they can harvest sufficient amount of energy for jamming. Furthermore, as in [48, 51], the channel between the transmitter and the AF relay is assumed to be perfectly known in a global fashion throughout the chapter. In addition, Alice, Bob and Eve are all equipped with single antenna, while the AF relay and each of the K helpers have the same number of N_t multiple antennas.

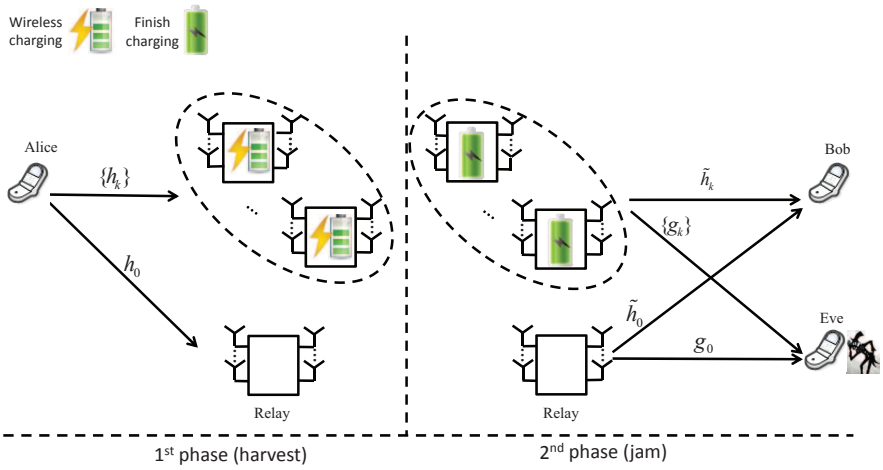
Using two equal slots for the HJ relaying protocol, as shown in Fig. 4.1(b), for the first phase, Alice sends a confidential message to the relay while simultaneously transferring energy to the K helpers; for the second phase, the relay amplifies and forwards the message to Bob while the K helpers perform CJ using their respective harvested energy from the first transmission phase, to compromise Eve. In this chapter, we assume a quasi-static fading environment and for convenience denote $\boldsymbol{h}_0 \in \mathbb{C}^{N_t \times 1}$ as the complex channel from the transmitter to the relay and $\boldsymbol{h}_k \in \mathbb{C}^{N_t \times 1}$, $k = 1, \ldots, K$, as that from the transmitter to the kth helper; $\tilde{\boldsymbol{h}}_0$ as the transpose of the complex channel from the relay to Bob and $\tilde{\boldsymbol{h}}_k \in \mathbb{C}^{N_t \times 1}$, $k = 1, \ldots, K$, as that from H_k to Bob; $\boldsymbol{g}_0 \in \mathbb{C}^{N_t \times 1}$ and $\boldsymbol{g}_k \in \mathbb{C}^{N_t \times 1}$, $k = 1, \ldots, K$, as those from the relay and H_k to Eve, respectively.

In the first transmission phase, the baseband received signal at the AF relay can be expressed as

$$\boldsymbol{y}_r = \boldsymbol{h}_0 \sqrt{P_s} s + \boldsymbol{n}_r, \tag{4.1}$$



(a) AF-relaying wiretap channel with jamming.



(b) The *HJ* relaying protocol.

Figure 4.1: HJ-enabled cooperative relaying for secure SWIPT.

where s is a circularly symmetric complex Gaussian (CSCG) RV, denoted by $s \sim \mathcal{CN}(0,1)$ and $\boldsymbol{n}_r$ is the additive complex noise vector, denoted by $\boldsymbol{n}_r \sim \mathcal{CN}(\boldsymbol{0}, \sigma_r^2 \boldsymbol{I})$. Also, P_s denotes the given transmit power at Alice. Further, the received signal at each helper H_k is expressed as

$$\boldsymbol{y}_k = \boldsymbol{h}_k \sqrt{P_s} s + \boldsymbol{n}'_k, \tag{4.2}$$

where $\boldsymbol{n}'_k$ is the additive noise, denoted by $\boldsymbol{n}'_k \sim \mathcal{CN}(\boldsymbol{0}, \sigma_h^2 \boldsymbol{I})$.

On the other hand, for WET, the harvested energy of H_k in each unit slot is given by

$$E_k = \eta \mathbb{E}[\|\boldsymbol{h}_k \sqrt{P_s} s\|^2] = \eta P_s \left\|\boldsymbol{h}_k\right\|^2, \ \forall k, \tag{4.3}$$

where $0 < \eta \leq 1$ denotes the EH efficiency.

In the second transmission phase, the linear operation at the AF relay can be represented by

$$\boldsymbol{x}' = \boldsymbol{W} \boldsymbol{y}_r, \tag{4.4}$$

where $\boldsymbol{x}' \in \mathbb{C}^{N_t \times 1}$ is the retransmit signal at the AF relay and $\boldsymbol{W} \in \mathbb{C}^{N_t \times N_t}$ is the beamforming matrix. Note that the transmit power of the AF relay can be shown as

$$\mathrm{tr}\left(\mathbb{E}\left[\boldsymbol{x}\boldsymbol{x}^H\right]\right) = \mathrm{tr}\left(\boldsymbol{W}\left(P_s \boldsymbol{h}_0 \boldsymbol{h}_0^H + \sigma_r^2 \boldsymbol{I}\right) \boldsymbol{W}^H\right), \tag{4.5}$$

which is constrained by the maximum available power at the AF relay, i.e., P_r, which is given by

$$\mathrm{tr}\left(\boldsymbol{W}\left(P_s \boldsymbol{h}_0 \boldsymbol{h}_0^H + \sigma_r^2 \boldsymbol{I}\right) \boldsymbol{W}^H\right) \leq P_r. \tag{4.6}$$

In the meantime, each H_k will help generate an AN $\boldsymbol{n}_k \in \mathbb{C}^{N_t \times 1}$ to interfere with Eve. Similar to [21], we assume that $\boldsymbol{n}_k$'s are independent CSCG vectors denoted by $\boldsymbol{n}_k \sim \mathcal{CN}(0, \boldsymbol{Q}_k)$, $\forall k$, since the worst-case noise for Eve is known to be Gaussian. In addition, each H_k has a transmit power constraint due to its harvested energy in the previous transmission phase, i.e., $\mathrm{tr}(\boldsymbol{Q}_k) \leq \eta P_s \|\boldsymbol{h}_k\|^2$ (c.f. (4.3)), $\forall k$.

The received signal at Bob can thus be expressed as

$$\boldsymbol{y}_b = \sqrt{P_s}\widetilde{\boldsymbol{h}}_0^T \boldsymbol{W}\boldsymbol{h}_0 s + \sum_{k=1}^{K} \widetilde{\boldsymbol{h}}_k^T \boldsymbol{n}_k + \widetilde{\boldsymbol{h}}_0^T \boldsymbol{W}\boldsymbol{n}_r + \boldsymbol{n}_b, \tag{4.7}$$

where $\boldsymbol{n}_b \sim \mathcal{CN}(0, \sigma_b^2 \boldsymbol{I})$ is the additive noise at Bob. Similarly, the received signal at Eve can be expressed as

$$\mathbf{y}_e = \sqrt{P_s}\boldsymbol{g}_0^T \boldsymbol{W}\boldsymbol{h}_0 s + \sum_{k=1}^{K} \boldsymbol{g}_k^T \boldsymbol{n}_k + \boldsymbol{g}_0^T \boldsymbol{W}\boldsymbol{n}_r + \boldsymbol{n}_e, \tag{4.8}$$

where $\boldsymbol{n}_e \sim \mathcal{CN}(0, \sigma_e^2 \boldsymbol{I})$. According to (4.7) and (4.8), the SINR at Bob and Eve can be, respectively, expressed as

$$\gamma_b = \frac{P_s |\widetilde{\boldsymbol{h}}_0^T \boldsymbol{W}\boldsymbol{h}_0|^2}{\sigma_r^2 \widetilde{\boldsymbol{h}}_0^T \boldsymbol{W}\boldsymbol{W}^H \widetilde{\boldsymbol{h}}_0^\dagger + \sum_{k=1}^{K} \widetilde{\boldsymbol{h}}_k^T \boldsymbol{Q}_k \widetilde{\boldsymbol{h}}_k^\dagger + \sigma_b^2}, \tag{4.9}$$

and

$$\gamma_e = \frac{P_s |\boldsymbol{g}_0^T \boldsymbol{W}\boldsymbol{h}_0|^2}{\sigma_r^2 \boldsymbol{g}_0^T \boldsymbol{W}\boldsymbol{W}^H \boldsymbol{g}_0^\dagger + \sum_{k=1}^{K} \boldsymbol{g}_k^T \boldsymbol{Q}_k \boldsymbol{g}_k^\dagger + \sigma_e^2}. \tag{4.10}$$

As such, the achievable secrecy rate at Bob is [21]

$$r_0 = \frac{1}{2}\left[\log_2(1+\gamma_b) - \log_2(1+\gamma_e)\right]^+. \tag{4.11}$$

4.4 Joint AN-AF Beamforming with Perfect CSI

4.4.1 Problem Formulation for Perfect CSI

We aim to maximize the secrecy rate at Bob subject to the transmit power constraints at the AF relay and each individual helper H_k, $k = 1, \ldots, K$. Thus, our

problem is to solve

$$\begin{aligned}(\text{P1}): \quad \max_{\{\boldsymbol{Q}_k\},\boldsymbol{W}} \quad & r_0 \\ \texttt{s.t.} \quad & (4.6), && \text{(4.12a)} \\ & \operatorname{tr}(\boldsymbol{Q}_k) \le \eta P_s \|\boldsymbol{h}_k\|^2, \ \forall k, && \text{(4.12b)} \\ & \boldsymbol{Q}_k \succeq \boldsymbol{0}, \ \forall k. && \text{(4.12c)}\end{aligned}$$

Next, we define a new function $\bar{F}(\{\boldsymbol{Q}_k\}, \boldsymbol{W})$ as

$$\bar{F}(\{\boldsymbol{Q}_k\}, \boldsymbol{W}) \triangleq \frac{1+\gamma_b}{1+\gamma_e}. \tag{4.13}$$

It can be easily shown that the optimal solution $\{\boldsymbol{Q}_k^*\}$, $\boldsymbol{W}^*$ to (P1), is also optimal for (P1′) given by

$$(\text{P1}'): \quad \max_{\{\boldsymbol{Q}_k\},\boldsymbol{W}} \bar{F}(\{\boldsymbol{Q}_k\}, \boldsymbol{W}) \ \texttt{s.t.} \ (4.12\text{a}) - (4.12\text{c}). \tag{4.14}$$

Hence, we focus on solving problem (P1′) in the rest of the chapter. However, since (P1′) is in general a non-convex problem that is hard to solve, we will reformulate it into a two-stage optimization problem. First, we constrain the SINR at Eve to be $\bar{\gamma}_e$, it thus follows from (4.13) that $\bar{F}(\{\overline{\boldsymbol{Q}}_k\}, \boldsymbol{W})$ is maximized when γ_b is maximized, which can be obtained by solving the following problem:

$$\begin{aligned}(\text{P1}'.1): \quad \max_{\{\boldsymbol{Q}_k,\boldsymbol{W}\}} \quad & \frac{P_s|\widetilde{\boldsymbol{h}}_0^T \boldsymbol{W} \boldsymbol{h}_0|^2}{\sigma_r^2 \widetilde{\boldsymbol{h}}_0^T \boldsymbol{W}\boldsymbol{W}^H \widetilde{\boldsymbol{h}}_0^\dagger + \sum_{k=1}^K \tilde{\boldsymbol{h}}_k^T \boldsymbol{Q}_k \tilde{\boldsymbol{h}}_k^\dagger + \sigma_b^2} && \text{(4.15)} \\ \texttt{s.t.} \quad & \frac{P_s|\boldsymbol{g}_0^T \boldsymbol{W} \boldsymbol{h}_0|^2}{\sigma_r^2 \boldsymbol{g}_0^T \boldsymbol{W}\boldsymbol{W}^H \boldsymbol{g}_0^\dagger + \sum_{k=1}^K \boldsymbol{g}_k^T \boldsymbol{Q}_k \boldsymbol{g}_k^\dagger + \sigma_e^2} = \bar{\gamma}_e, && \text{(4.16)} \\ & (4.12\text{a}) - (4.12\text{c}). && \text{(4.17)}\end{aligned}$$

Let $H(\bar{\gamma}_e)$ denote the optimal value of (P1′.1) given $\bar{\gamma}_e$. Then (P1′) can be equivalently

solved by

$$(\text{P1}'.2): \max_{\bar{\gamma}_e>0} \ \frac{1+H(\bar{\gamma}_e)}{1+\bar{\gamma}_e}. \tag{4.18}$$

Lemma 4.4.1. *Problem* (P1′) *has the same optimal value as* (P1′.2)*, and the same optimal solution as* (P1′.1) *when* $\bar{\gamma}_e$ *takes the optimal solution for* (P1′.2).

Proof. The proof follows from [57, *Lemmas 4.1-4.2*]. □

Therefore, (P1′) can be solved in the following two steps. First, given any $\bar{\gamma}_e > 0$, we solve (P1′.1) to attain $H(\bar{\gamma}_e)$; then we solve (P1′.2) to obtain the optimal $\bar{\gamma}_e^*$.

4.4.2 Optimal Solution for Perfect CSI

Here, we consider solving problem (P1′.1) by jointly optimizing the covariance matrix for the AN at each of the HJ helper, $\boldsymbol{Q}_k$'s, and the beamforming matrix, $\boldsymbol{W}$. To facilitate the analysis in the sequel, we rewrite the following equations in line with our definition of vec(·) [91, Chapter 13]:

$$|\tilde{\boldsymbol{h}}_0^T \boldsymbol{W} \boldsymbol{h}_0|^2 = |\text{vec}^T(\tilde{\boldsymbol{h}}_0 \boldsymbol{h}_0^T)\text{vec}(\boldsymbol{W})|^2, \tag{4.19}$$

$$\tilde{\boldsymbol{h}}_0^T \boldsymbol{W} \boldsymbol{W}^H \tilde{\boldsymbol{h}}_0^\dagger = \|(\tilde{\boldsymbol{h}}_0^T \otimes \boldsymbol{I})\text{vec}(\boldsymbol{W})\|^2, \tag{4.20}$$

$$\boldsymbol{g}_0^T \boldsymbol{W} \boldsymbol{W}^H \boldsymbol{g}_0^\dagger = \|(\boldsymbol{g}_0^T \otimes \boldsymbol{I})\text{vec}(\boldsymbol{W})\|^2. \tag{4.21}$$

In addition, $\text{tr}(\boldsymbol{W}(P_s \boldsymbol{h}_0 \boldsymbol{h}_0^H + \sigma_r^2 \boldsymbol{I})\boldsymbol{W}^H) = \|\boldsymbol{\Phi}\boldsymbol{w}\|^2$, where $\boldsymbol{\Phi} = (\boldsymbol{I} \otimes \boldsymbol{\Theta}^T)^{1/2}$ with $\boldsymbol{\Theta} = P_s \boldsymbol{h}_0 \boldsymbol{h}_0^H + \sigma_r^2 \boldsymbol{I}$. Hence, (P1′.1) can be rewritten as

(P1′.1-RW):

$$\max_{\boldsymbol{W}, \{\boldsymbol{Q}_k\}} \ \frac{P_s |\boldsymbol{f}_1^T \boldsymbol{w}|^2}{\sigma_r^2 \|\boldsymbol{Y}_1 \boldsymbol{w}\|^2 + \sum_{k=1}^K \tilde{\boldsymbol{h}}_k^T \boldsymbol{Q}_k \tilde{\boldsymbol{h}}_k^\dagger + \sigma_b^2}$$

$$\text{s.t.} \ \frac{P_s |\boldsymbol{f}_2^T \boldsymbol{w}|^2}{\sigma_r^2 \|\boldsymbol{Y}_2 \boldsymbol{w}\|^2 + \sum_{k=1}^K \boldsymbol{g}_k^T \boldsymbol{Q}_k \boldsymbol{g}_k^\dagger + \sigma_e^2} = \bar{\gamma}_e, \tag{4.22a}$$

$$\|\Phi \boldsymbol{w}\|^2 \le P_r, \tag{4.22b}$$

$$(4.12\text{b}), (4.12\text{c}),$$

in which $\boldsymbol{w} = \text{vec}(\boldsymbol{W})$, $\boldsymbol{f}_1 = \text{vec}(\widetilde{\boldsymbol{h}}_0 \boldsymbol{h}_0^T)$, $\boldsymbol{f}_2 = \text{vec}(\boldsymbol{g}_0 \boldsymbol{h}_0^T)$, $\boldsymbol{Y}_1 = \widetilde{\boldsymbol{h}}_0^T \otimes \boldsymbol{I}$ and $\boldsymbol{Y}_2 = \boldsymbol{g}_0^T \otimes \boldsymbol{I}$.

As problem (P1′.1-RW) is non-convex, we define $\boldsymbol{X} \triangleq \boldsymbol{w}\boldsymbol{w}^H$, $\boldsymbol{F}_1 \triangleq \boldsymbol{f}_1^\dagger \boldsymbol{f}_1^T$, $\boldsymbol{F}_2 \triangleq \boldsymbol{f}_2^\dagger \boldsymbol{f}_2^T$, $\overline{\boldsymbol{Y}}_1 \triangleq \boldsymbol{Y}_1^H \boldsymbol{Y}_1$, $\overline{\boldsymbol{Y}}_2 \triangleq \boldsymbol{Y}_2^H \boldsymbol{Y}_2$ and $\overline{\boldsymbol{\Phi}} \triangleq \boldsymbol{\Phi}^H \boldsymbol{\Phi}$. Then by ignoring the rank-one constraint on $\boldsymbol{X}$, (P1′.1-RW) is modified as

$$
\begin{aligned}
&\text{(P1′.1-RW-SDR-Eqv)}: \\
&\max_{\boldsymbol{X}, \{\boldsymbol{Q}_k\}} \quad \frac{P_s \text{tr}(\boldsymbol{F}_1 \boldsymbol{X})}{\sigma_r^2 \text{tr}(\overline{\boldsymbol{Y}}_1 \boldsymbol{X}) + \sum_{k=1}^K \tilde{\boldsymbol{h}}_k^T \boldsymbol{Q}_k \tilde{\boldsymbol{h}}_k^\dagger + \sigma_b^2} \\
&\text{s.t.} \quad P_s \text{tr}(\boldsymbol{F}_2 \boldsymbol{X}) \\
&\qquad = \bar{\gamma}_e \left(\sigma_r^2 \text{tr}(\overline{\boldsymbol{Y}}_2 \boldsymbol{X}) + \sum_{k=1}^K \boldsymbol{g}_k^T \boldsymbol{Q}_k \boldsymbol{g}_k^\dagger + \sigma_e^2 \right), \qquad (4.23a) \\
&\qquad \text{tr}(\overline{\boldsymbol{\Phi}} \boldsymbol{X}) \leq P_r, \qquad (4.23b) \\
&\qquad \text{tr}\left(\boldsymbol{Q}_k\right) \leq \eta P_s \left\|\boldsymbol{h}_k\right\|^2, \ \forall k, \qquad (4.23c) \\
&\qquad \boldsymbol{X} \succeq \boldsymbol{0}, \boldsymbol{Q}_k \succeq \boldsymbol{0}, \ \forall k. \qquad (4.23d)
\end{aligned}
$$

Problem (P1′.1-RW-SDR-Eqv), via Charnes-Cooper transformation [92], can be equivalently recast as

$$
\begin{aligned}
&\text{(P1′.1-RW-SDR)}: \max_{\boldsymbol{X}, \{\boldsymbol{Q}_k\}, \tau} \quad P_s \text{tr}(\boldsymbol{F}_1 \boldsymbol{X}) \\
&\text{s.t.} \quad \sigma_r^2 \text{tr}(\overline{\boldsymbol{Y}}_1 \boldsymbol{X}) + \sum_{k=1}^K \tilde{\boldsymbol{h}}_k^T \boldsymbol{Q}_k \tilde{\boldsymbol{h}}_k^\dagger + \tau \sigma_b^2 = 1, \qquad (4.24a) \\
&P_s \text{tr}(\boldsymbol{F}_2 \boldsymbol{X}) = \bar{\gamma}_e \left(\sigma_r^2 \text{tr}(\overline{\boldsymbol{Y}}_2 \boldsymbol{X}) + \sum_{k=1}^K \boldsymbol{g}_k^T \boldsymbol{Q}_k \boldsymbol{g}_k^\dagger + \tau \sigma_e^2 \right), \qquad (4.24b) \\
&\text{tr}(\overline{\boldsymbol{\Phi}} \boldsymbol{X}) \leq \tau P_r, \qquad (4.24c) \\
&\text{tr}\left(\boldsymbol{Q}_k\right) \leq \tau \eta P_s \left\|\boldsymbol{h}_k\right\|^2, \ \forall k, \qquad (4.24d) \\
&\boldsymbol{X} \succeq \boldsymbol{0}, \boldsymbol{Q}_k \succeq \boldsymbol{0}, \ \forall k, \tau \geq 0. \qquad (4.24e)
\end{aligned}
$$

Lemma 4.4.2. *The constraints in* (4.24a) *and* (4.24b) *can be replaced by*

$\sigma_r^2\mathrm{tr}(\overline{\boldsymbol{Y}}_1\boldsymbol{X}) + \sum_{k=1}^{K}\tilde{\boldsymbol{h}}_k^T\boldsymbol{Q}_k\tilde{\boldsymbol{h}}_k^\dagger + \tau\sigma_b^2 \leq 1$ *and* $P_s\mathrm{tr}(\boldsymbol{F}_2\boldsymbol{X}) \leq \bar{\gamma}_e(\sigma_r^2\mathrm{tr}(\overline{\boldsymbol{Y}}_2\boldsymbol{X}) + \sum_{k=1}^{K}\boldsymbol{g}_k^T\boldsymbol{Q}_k\boldsymbol{g}_k^\dagger + \tau\sigma_e^2)$*, respectively, where both inequalities will be activated when problem* (P1′) *obtains its optimum value.*

Proof. See [93, Appendix A]. □

Since problem (P1′.1-RW-SDR) is a standard convex optimization problem and satisfies the Slater's condition, its gap with its dual problem is zero [84]. Now, let λ denote the dual variable associated with the equality constraint in (4.24a), α associated with the other equality constraint in (4.24b), β_0 associated with the transmit power constraint for the AF relay in (4.24c), $\{\beta_k\}$ associated with the transmit power constraints for each H_k in (4.24d), and ζ associated with τ. Then the Lagrangian of problem (P1′.1-RW-SDR) is given by

$$L(\boldsymbol{\Omega}) = \mathrm{tr}(\boldsymbol{A}\boldsymbol{X}) + \sum_{k=1}^{K}\mathrm{tr}(\boldsymbol{B}_k\boldsymbol{Q}_k) + \zeta\tau + \lambda, \tag{4.25}$$

where $\boldsymbol{\Omega}$ denotes the set of all primal and dual variables,

$$\boldsymbol{A} = P_s\boldsymbol{F}_1 - \lambda\sigma_r^2\overline{\boldsymbol{Y}}_1 - \alpha P_s\boldsymbol{F}_2 + \alpha\bar{\gamma}_e\sigma_r^2\overline{\boldsymbol{Y}}_2 - \beta_0\overline{\boldsymbol{\Phi}}, \tag{4.26}$$

$$\boldsymbol{B}_k = -\lambda\tilde{\boldsymbol{h}}_k^*\tilde{\boldsymbol{h}}_k^T + \alpha\bar{\gamma}_e\boldsymbol{g}_k^*\boldsymbol{g}_k^T - \beta_k\boldsymbol{I},\ \forall k, \tag{4.27}$$

$$\zeta = -\lambda\sigma_b^2 + \alpha\bar{\gamma}_e\sigma_e^2 + \beta_0 P_r + \sum_{k=1}^{K}\eta P_s\beta_k\|\boldsymbol{h}_k\|^2. \tag{4.28}$$

Proposition 4.4.1. *The optimal solution,* $(\boldsymbol{X}^*, \{\boldsymbol{Q}_k^*\}, \tau^*)$*, to* (P1′.1-RW-SDR) *satisfies the following conditions:*

1. $\mathrm{rank}(\boldsymbol{Q}_k)\begin{cases} \geq N_t - 2, & \text{if } \beta_k^* = 0, \\ = 1, & \text{if } \beta_k^* > 0, \end{cases} \forall k;$

2. $\boldsymbol{X}^*$ *can be expressed as*

$$\boldsymbol{X}^* = \sum_{n=1}^{N_t^2 - r_c} a_n\boldsymbol{\eta}_n\boldsymbol{\eta}_n^H + b\boldsymbol{\xi}\boldsymbol{\xi}^H, \tag{4.29}$$

where $a_n \geq 0\ \forall n$, $b > 0$, $r_c = \mathrm{rank}(\boldsymbol{C}^*)$ *(c.f.* (C.2)*) and* $\boldsymbol{\xi} \in \mathbb{C}^{N_t^2 \times 1}$ *is a vector orthogonal to* $\boldsymbol{\Xi} = \{\boldsymbol{\eta}_n\}_{n=1}^{N_t^2 - r_c}$*, which consists of orthonormal basis for* $\mathbf{null}(\boldsymbol{C}^*)$*;*

3. *According to* (4.29)*, if* $\mathrm{rank}(\boldsymbol{X}^*) > 1$*, then we have the following sufficient condition to yield an optimal solution of* $\boldsymbol{X}$ *with rank-one:*

$$\hat{\boldsymbol{X}}^* = b\boldsymbol{\xi}\boldsymbol{\xi}^H, \tag{4.30}$$

$$\hat{\boldsymbol{Q}}_k^* = \boldsymbol{Q}_k^*,\ \forall k, \tag{4.31}$$

$$\hat{\tau}^* = \tau^* + \Delta\tau, \tag{4.32}$$

is also optimal to problem (P1′.1-RW-SDR)*, if there exists* $\Delta\tau \geq 0$ *such that*

$$\left[\sum_{n=1}^{N_t^2 - r_c} a_n \mathrm{tr}\left(\boldsymbol{\eta}_n^H \left(\frac{\sigma_r^2 \overline{\boldsymbol{Y}}_2}{\sigma_e^2} - \frac{P_s \boldsymbol{F}_2}{\gamma_e \sigma_e^2}\right)\boldsymbol{\eta}_n\right)\right]^+ \leq \Delta\tau \leq \frac{\sigma_r^2}{\sigma_b^2} \sum_{n=1}^{N_t^2 - r_c} a_n \mathrm{tr}\left(\boldsymbol{\eta}_n^H \overline{\boldsymbol{Y}}_1 \boldsymbol{\eta}_n\right). \tag{4.33}$$

Proof. See Appendix C. □

Note from Proposition 4.4.1 that if $\mathrm{rank}(\boldsymbol{X}^*) = 1$, then the optimal $\boldsymbol{w}^*$ to (P1′.1-RW) can be found directly from the eigenvalue decomposition (EVD) of $\overline{\boldsymbol{X}}^*$, where $\overline{\boldsymbol{X}}^* = \boldsymbol{X}^*/\tau^*$. Namely, the upper-bound optimum value obtained by solving (P1′.1-RW-SDR) is tight in this case; otherwise, $(\boldsymbol{X}^*, \{\boldsymbol{Q}_k^*\}, \tau^*)$ only serves as an upper-bound solution.

Now, we show that this upper-bound is always achievable by a rank-one $\boldsymbol{X}^*$. When $\mathrm{rank}(\boldsymbol{X}^*) > 1$, firstly, we check whether the sufficient condition proposed in (4.33) is satisfied. If it is met, then a direct reconstruction of $(\hat{\boldsymbol{X}}^*, \{\hat{\boldsymbol{Q}}_k^*\}, \hat{\tau}^*)$ with $\mathrm{rank}(\hat{\boldsymbol{X}}^*) = 1$ follows according to (4.30)–(4.32); otherwise, assume that any optimal solution to problem (P1′.1-RW-SDR) has no zero component, i.e., $(\boldsymbol{X}^* \neq \boldsymbol{0}, \{\boldsymbol{Q}_k^* \neq \boldsymbol{0}\}, \tau^* \neq 0)$. In addition, the number of optimization variables and the number of shaping constraints are denoted by L and M, respectively. Since $L = K + 2$ and $M = K + 3$ for (P1′.1-RW-SDR), we have $M \leq L + 2$ satisfied. Thus, according

to [94, *Proposition 3.5*], (P1′.1-RW-SDR) has an optimal solution of $\hat{\boldsymbol{X}}^*$ that is rank-one. Also, the detailed rank reduction procedure based on an arbitrary-rank solution has been given in [94, Algorithm 1]. Algorithm 4.1 for solving (P1′) is shown in Table 5.1.

Table 4.1: Algorithm I for (P1′)

- **Initialize** $\bar{\gamma}_{e\_\text{search}} = 0 : \alpha : \bar{\gamma}_{e\,\max}$ and $i = 0$
- **Repeat**
  1) **Set** $i = i + 1$;
  2) Given $\bar{\gamma}_e = \bar{\gamma}_{e\_\text{search}}(i)$,
  solve (P1′.1-RW-SDR) and **obtain** $H(\bar{\gamma}_e^{(i)})$.
- **Until** $i = L$, where $L = \lfloor \frac{\bar{\gamma}_{e\,\max}}{\alpha} \rfloor + 1$ is the length of $\bar{\gamma}_{e\_\text{search}}$
- **Set** $\bar{\gamma}_e^* = \bar{\gamma}_{e\_\text{search}} \left(\arg\max_i \left\{ \frac{1+H(\bar{\gamma}_e^{(i)})}{1+\bar{\gamma}_e^{(i)}} \right\} \right)$ for (P1′.2)
- Given $\bar{\gamma}_e^*$, **solve** (P1′.1-RW-SDR) to obtain $(\boldsymbol{X}^*, \{\boldsymbol{Q}_k^*\}, \tau^*)$
  if $\text{rank}(\boldsymbol{X}^*) = 1$, **apply** EVD on $\boldsymbol{X}^*$ such that $\boldsymbol{X}^* = \boldsymbol{w}^* \boldsymbol{w}^{*H}$;
  else if the sufficient condition in (4.33) is satisfied,
  construct $(\hat{\boldsymbol{X}}^*, \{\hat{\boldsymbol{Q}}_k^*\}, \hat{\tau}^*)$ following (4.30)-(4.32) and **set** $\boldsymbol{w}^* = \sqrt{b}\boldsymbol{\xi}$;
  else construct $\hat{\boldsymbol{X}}^*$ using the procedure in [94, Algorithm 1].
  end
  end
- **Recover** $\boldsymbol{W}^* = \text{vec}^{-1}(\boldsymbol{w}^*)$

4.4.3 Suboptimal Solutions for Perfect CSI

Optimal Solution Structure based Scheme

We propose a relay beamforming design for (P1′.1) based on the optimal structure of $\boldsymbol{W}$ [10, Theorem 3.1]. First, define $\boldsymbol{H}_1 \triangleq [\tilde{\boldsymbol{h}}_0 \ \boldsymbol{g}_0]$ and $\boldsymbol{H}_2 \triangleq [\boldsymbol{h}_0 \ \boldsymbol{g}_0]$. Then express the truncated singular-value decomposition (SVD) of $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$,

respectively, as

$$\boldsymbol{H}_1 = \boldsymbol{U}_1\boldsymbol{\Sigma}_1\boldsymbol{V}_1^H, \tag{4.34}$$

$$\boldsymbol{H}_2 = \boldsymbol{U}_2\boldsymbol{\Sigma}_2\boldsymbol{V}_2^H. \tag{4.35}$$

Lemma 4.4.3. *The optimal relay beamforming matrix $\boldsymbol{W}$ for problem* (P1′.1) *is of the form:*

$$\boldsymbol{W} = \boldsymbol{U}_1^\dagger\boldsymbol{B}\boldsymbol{U}_2^H + \boldsymbol{U}_1^\dagger\boldsymbol{C}(\boldsymbol{U}_2^\perp)^H, \tag{4.36}$$

where $\boldsymbol{B} \in \mathbb{C}^{2\times 2}$ and $\boldsymbol{C} \in \mathbb{C}^{2\times(N_t-2)}$ are two unknown matrices, and $\boldsymbol{U}_1^\perp$, $\boldsymbol{U}_2^\perp \in \mathbb{C}^{N_t\times(N_t-2)}$ satisfy $\boldsymbol{U}_1^\perp(\boldsymbol{U}_1^\perp)^H = \boldsymbol{I} - \boldsymbol{U}_1\boldsymbol{U}_1^H$, $\boldsymbol{U}_2^\perp(\boldsymbol{U}_2^\perp)^H = \boldsymbol{I} - \boldsymbol{U}_2\boldsymbol{U}_2^H$, respectively.

Proof. First, we construct $\boldsymbol{W}$ as

$$\begin{aligned}\boldsymbol{W} &= [\boldsymbol{U}_1^\dagger, (\boldsymbol{U}_1^\perp)^\dagger]\begin{bmatrix}\boldsymbol{B} & \boldsymbol{C}\\ \boldsymbol{D} & \boldsymbol{E}\end{bmatrix}[\boldsymbol{U}_2, \boldsymbol{U}_2^\perp]^H \\ &= \boldsymbol{U}_1^\dagger\boldsymbol{B}\boldsymbol{U}_2^H + \boldsymbol{U}_1^\dagger\boldsymbol{C}(\boldsymbol{U}_2^\perp)^H + (\boldsymbol{U}_1^\perp)^\dagger\boldsymbol{D}\boldsymbol{U}_2^H + (\boldsymbol{U}_1^\perp)^\dagger\boldsymbol{E}(\boldsymbol{U}_2^\perp)^H,\end{aligned} \tag{4.37}$$

where $\boldsymbol{B} \in \mathbb{C}^{2\times 2}$, $\boldsymbol{C} \in \mathbb{C}^{2\times(N_t-2)}$, $\boldsymbol{D} \in \mathbb{C}^{(N_t-2)\times 2}$ and $\boldsymbol{E} \in \mathbb{C}^{(N_t-2)\times(N_t-2)}$ are undetermined matrices. Then according to (4.34) and (4.35), it follows that $|\widetilde{\boldsymbol{h}}_0^T\boldsymbol{W}\boldsymbol{h}_0|^2 = |\widetilde{\boldsymbol{h}}_0^T\boldsymbol{U}_1^\dagger\boldsymbol{B}\boldsymbol{U}_2^H\boldsymbol{h}_0|^2$ and $\widetilde{\boldsymbol{h}}_0^T\boldsymbol{W}\boldsymbol{W}^H\widetilde{\boldsymbol{h}}_0^\dagger = \|\boldsymbol{B}^H\boldsymbol{U}_1^T\widetilde{\boldsymbol{h}}_0^\dagger\|^2 + \|\boldsymbol{C}^H\boldsymbol{U}_1^T\widetilde{\boldsymbol{h}}_0^\dagger\|^2$. Similarly, we also have $|\boldsymbol{g}_0^T\boldsymbol{W}\boldsymbol{h}_0|^2 = |\boldsymbol{g}_0^T\boldsymbol{U}_1^\dagger\boldsymbol{B}\boldsymbol{U}_2^H\boldsymbol{h}_0|^2$ and $\boldsymbol{g}_0^T\boldsymbol{W}\boldsymbol{W}^H\boldsymbol{g}_0^\dagger = \|\boldsymbol{g}_0^T\boldsymbol{U}_1^\dagger\boldsymbol{B}\|^2 + \|\boldsymbol{g}_0^T\boldsymbol{U}_1^\dagger\boldsymbol{C}\|^2$. Thus, γ_b (c.f. (4.9)) and γ_e (c.f. (4.10)) do not depend on $\boldsymbol{D}$ and $\boldsymbol{E}$.

Next, by substituting (4.37) for $\boldsymbol{W}$ in (4.5), we have $P_r \geq P_s(\|\boldsymbol{B}\boldsymbol{U}_2^H\boldsymbol{h}_0\|^2 + \|\boldsymbol{D}\boldsymbol{U}_2^H\boldsymbol{h}_0\|^2) + \sigma_r^2\mathrm{tr}(\boldsymbol{B}^H\boldsymbol{B} + \boldsymbol{C}^H\boldsymbol{C} + \boldsymbol{D}^H\boldsymbol{D} + \boldsymbol{E}^H\boldsymbol{E})$. Since (P1′) is a secrecy rate maximization problem subject to the given P_r, it turns out that given the optimum secrecy rate, P_r is the minimized required power by taking $\boldsymbol{D} = \boldsymbol{0}$ and $\boldsymbol{E} = \boldsymbol{0}$, while $\boldsymbol{B}$ and $\boldsymbol{C}$ cannot be determined directly. Thus, $\boldsymbol{W} = \boldsymbol{U}_1^\dagger\boldsymbol{B}\boldsymbol{U}_2^H + \boldsymbol{U}_1^\dagger\boldsymbol{C}(\boldsymbol{U}_2^\perp)^H$. □

Denote $\boldsymbol{U}_1^H\tilde{\boldsymbol{h}}_0$, $\boldsymbol{U}_2^H\boldsymbol{h}_0$, $\boldsymbol{U}_1^H\boldsymbol{g}_0$ by $\bar{\tilde{\boldsymbol{h}}}_0$, $\bar{\boldsymbol{h}}_0$, $\bar{\boldsymbol{g}}_0$, respectively. We thus simplify $|\widetilde{\boldsymbol{h}}_0^T\boldsymbol{W}\boldsymbol{h}_0|^2$ and $|\boldsymbol{g}_0^T\boldsymbol{W}\boldsymbol{h}_0|^2$ as $|\bar{\tilde{\boldsymbol{h}}}_0^T\boldsymbol{B}\bar{\boldsymbol{h}}_0|^2$ and $|\bar{\boldsymbol{g}}_0^T\boldsymbol{B}\bar{\boldsymbol{h}}_0|^2$, respectively. Since $\boldsymbol{C}$ has

$2(N_t-2)$ complex variables, we devise a suboptimal design for $\boldsymbol{C}$ to reduce the size of variables by (N_t-2). Specifically, let $\boldsymbol{C}=\boldsymbol{u}'^{\perp}\boldsymbol{v}^T$, where $\boldsymbol{u}'=\bar{\tilde{\boldsymbol{h}}}_0^{\dagger}/\|\bar{\tilde{\boldsymbol{h}}}_0\|$ such that $\boldsymbol{u}'^{\perp}\boldsymbol{u}'^{\perp H}=\boldsymbol{I}-\boldsymbol{u}'\boldsymbol{u}'^{H}$. Hence, $\widetilde{\boldsymbol{h}}_0^T\boldsymbol{W}\boldsymbol{W}^H\widetilde{\boldsymbol{h}}_0^{\dagger}$, $\boldsymbol{g}_0^T\boldsymbol{W}\boldsymbol{W}^H\boldsymbol{g}_0^{\dagger}$ and (4.5) can be reduced to $\|\boldsymbol{B}^H\bar{\tilde{\boldsymbol{h}}}_0^{\dagger}\|^2$, $\|\boldsymbol{B}^H\bar{\boldsymbol{g}}_0^{\dagger}\|^2+|\boldsymbol{v}^{\dagger}\boldsymbol{u}'^{\perp H}\bar{\boldsymbol{g}}_0^{\dagger}|^2$ and $P_s\|\boldsymbol{B}\bar{\boldsymbol{h}}_0\|^2+\sigma_r^2\mathrm{tr}(\boldsymbol{B}^H\boldsymbol{B})+\sigma_r^2\|\boldsymbol{v}\|^2$, respectively. Then define $\boldsymbol{b}=\mathrm{vec}(\boldsymbol{B})$, $\bar{\boldsymbol{f}}_1=\mathrm{vec}(\bar{\tilde{\boldsymbol{h}}}_0\bar{\boldsymbol{h}}_0^T)$, $\bar{\boldsymbol{f}}_2=\mathrm{vec}(\bar{\boldsymbol{g}}_0\bar{\boldsymbol{h}}_0^T)$, $\boldsymbol{Y}'_1=\bar{\tilde{\boldsymbol{h}}}_0^T\otimes\boldsymbol{I}$, $\boldsymbol{Y}'_2=\bar{\boldsymbol{g}}_0^T\otimes\boldsymbol{I}$, and $\boldsymbol{\Phi}'=(\boldsymbol{I}\otimes\boldsymbol{\Theta}'^T)^{1/2}$ with $\boldsymbol{\Theta}'=P_s\bar{\boldsymbol{h}}_0\bar{\boldsymbol{h}}_0^H+\sigma_r^2\boldsymbol{I}$; $\boldsymbol{Z}=\boldsymbol{b}\boldsymbol{b}^H$, $\boldsymbol{V}=\boldsymbol{v}\boldsymbol{v}^H$, $\overline{\boldsymbol{F}}_1=\bar{\boldsymbol{f}}_1^{\dagger}\bar{\boldsymbol{f}}_1^T$, $\overline{\boldsymbol{F}}_2=\bar{\boldsymbol{f}}_2^{\dagger}\bar{\boldsymbol{f}}_2^T$, $\overline{\boldsymbol{Y}}'_1=\boldsymbol{Y}'^H_1\boldsymbol{Y}'_1$, $\overline{\boldsymbol{Y}}'_2=\boldsymbol{Y}'^H_2\boldsymbol{Y}'_2$, and $\overline{\boldsymbol{\Phi}}'=\boldsymbol{\Phi}'^H\boldsymbol{\Phi}'$. The suboptimal design for problem (P1′.1) by ignoring the rank constraints on $\boldsymbol{Z}$ and $\boldsymbol{V}$ is thus given by

$$
\begin{aligned}
&(\text{P1}'.1\text{-sub1-SDR}): \max_{\boldsymbol{Z},\boldsymbol{V},\{\boldsymbol{Q}_k\},\tau} \quad P_s\mathrm{tr}(\overline{\boldsymbol{F}}_1\boldsymbol{Z})\\
&\mathtt{s.t.}\ \ \sigma_r^2\mathrm{tr}(\overline{\boldsymbol{Y}}'_1\boldsymbol{Z})+\sum_{k=1}^{K}\tilde{\boldsymbol{h}}_k^T\boldsymbol{Q}_k\tilde{\boldsymbol{h}}_k^{\dagger}+\tau\sigma_b^2=1, && (4.38a)\\
&P_s\mathrm{tr}(\overline{\boldsymbol{F}}_2\boldsymbol{Z})\le\bar{\gamma}_e\left(\sigma_r^2\left(\mathrm{tr}(\overline{\boldsymbol{Y}}'_2\boldsymbol{Z})+|\bar{\boldsymbol{g}}_0^T\boldsymbol{u}'^{\perp}|^2\mathrm{tr}(\boldsymbol{V})\right)+\sum_{k=1}^{K}\boldsymbol{g}_k^T\boldsymbol{Q}_k\boldsymbol{g}_k^{\dagger}+\tau\sigma_e^2\right), && (4.38b)\\
&\mathrm{tr}(\overline{\boldsymbol{\Phi}}'\boldsymbol{Z})+\sigma_r^2\mathrm{tr}(\boldsymbol{Z})+\sigma_r^2\mathrm{tr}(\boldsymbol{V})\le\tau P_r, && (4.38c)\\
&\mathrm{tr}\left(\boldsymbol{Q}_k\right)\le\tau\eta P_s\left\|\boldsymbol{h}_k\right\|^2,\ \forall k, && (4.38d)\\
&\tau\ge 0,\ \boldsymbol{Q}_k\succeq\boldsymbol{0},\ \forall k,\ \boldsymbol{Z}\succeq\boldsymbol{0},\ \boldsymbol{V}\succeq\boldsymbol{0}. && (4.38e)
\end{aligned}
$$

Remark 4.4.1. *The variables in* (P1′.1-sub1-SDR)*, i.e.,* $\boldsymbol{Z}\in\mathbb{C}^{4\times4}$*,* $\boldsymbol{V}\in\mathbb{C}^{(N_t-2)\times(N_t-2)}$*, are of much reduced size. Further, the reconstruction of* $\boldsymbol{v}^*$ *from* $\boldsymbol{V}$ *can be briefly explained as follows. Given the Lagrangian of* (P1′.1-sub1-SDR)*, the KKT conditions w.r.t.* $\boldsymbol{V}^*$ *are given by*

$$
\begin{aligned}
(\alpha^*\bar{\gamma}_e|\bar{\boldsymbol{g}}_0^T\boldsymbol{u}'^{\perp}|^2-\beta_0^*\sigma_r^2)\boldsymbol{I}+\boldsymbol{U}^*=0, && (4.39)\\
\boldsymbol{U}^*\boldsymbol{V}^*=0. && (4.40)
\end{aligned}
$$

Post-multiplying (4.39) with $\boldsymbol{V}^*$, we have $(\alpha^*\bar{\gamma}_e|\bar{\boldsymbol{g}}_0^T\boldsymbol{u}'^{\perp}|^2-\beta_0^*\sigma_r^2)\boldsymbol{V}^*=0$. As a result, if $\frac{\alpha^*}{\beta_0^*}\neq\frac{\sigma_r^2}{\bar{\gamma}_e|\bar{\boldsymbol{g}}_0^T\boldsymbol{u}'^{\perp}|^2}$, $\boldsymbol{V}^*=\boldsymbol{0}$; otherwise $\boldsymbol{V}^*=\boldsymbol{v}^*\boldsymbol{v}^{*H}$, with $\boldsymbol{v}^*=\sqrt{\mathrm{tr}(\boldsymbol{V}^*)}\boldsymbol{v}_0$,

where $\boldsymbol{v}_0 \in \mathbb{C}^{(N_t-2)\times 1}$ is an arbitrary vector with unit norm. With $\boldsymbol{V}$ solved, (P1′.1-sub1-SDR) reduces to a problem with similar structure as (P1′.1-RW-SDR), and the proof for existence of a rank-one $\boldsymbol{Z}$ can be referred to Proposition 4.4.1.

Zero-forcing

We propose a low-complexity zero-forcing (ZF) scheme for (P1′.1), in which the jamming signal places a null at Bob, and then a semi-closed form solution for $\boldsymbol{W}$ is derived. In line with the principle of ZF jamming [36], the jamming signal $\boldsymbol{n}_k$ is designed as $\boldsymbol{n}_k = \tilde{\boldsymbol{V}}_k\tilde{\boldsymbol{n}}_k$ such that $\boldsymbol{I} - \frac{\tilde{\boldsymbol{h}}_k^\dagger \tilde{\boldsymbol{h}}_k^T}{\|\tilde{\boldsymbol{h}}_k\|^2} = \tilde{\boldsymbol{V}}_k\tilde{\boldsymbol{V}}_k^H$, and $\tilde{\boldsymbol{n}}_k \in \mathbb{C}^{(N_t-1)\times 1}$ is an arbitrary random vector, $\tilde{\boldsymbol{n}}_k \sim \mathcal{CN}(\boldsymbol{0}, \tilde{\boldsymbol{Q}}_k)$, $k = 1, \ldots, K$. Thus, given any $\boldsymbol{W}$, $\tilde{\boldsymbol{n}}_k$'s can be optimized to maximize the effect of jamming at Eve by $\max_{\tilde{\boldsymbol{Q}}_k} \sum_{k=1}^K \boldsymbol{g}_k^T\tilde{\boldsymbol{V}}_k\tilde{\boldsymbol{Q}}_k\tilde{\boldsymbol{V}}_k^H\boldsymbol{g}_k^\dagger$, which gives $\tilde{\boldsymbol{Q}}_k^* = \zeta_k^2\tilde{\boldsymbol{g}}_k^\dagger\tilde{\boldsymbol{g}}_k^T$, where $\tilde{\boldsymbol{g}}_k = \tilde{\boldsymbol{V}}_k^T\boldsymbol{g}_k$, and $\zeta_k = \sqrt{\eta P_s}\|\boldsymbol{h}_k\|/\|\tilde{\boldsymbol{g}}_k\|$ is determined by (4.24d), $\forall k$. As such, $\sum_{k=1}^K \boldsymbol{g}_k^T\tilde{\boldsymbol{V}}_k\tilde{\boldsymbol{Q}}_k^*\tilde{\boldsymbol{V}}_k^H\boldsymbol{g}_k^\dagger$ turns out to be $\sum_{k=1}^K \eta P_s\|\boldsymbol{h}_k\|^2\|\tilde{\boldsymbol{g}}_k\|^2$, which is denoted by q.

With fixed q, (P1′.1-RW-SDR) can be recast as

$$
\begin{aligned}
&\text{(P1}'\text{.1-sub2-SDR)}: \ \max_{\boldsymbol{X},\tau} \ P_s\mathrm{tr}(\boldsymbol{F}_1\boldsymbol{X}) \\
&\text{s.t. } \sigma_r^2\mathrm{tr}(\overline{\boldsymbol{Y}}_1\boldsymbol{X}) + \tau\sigma_b^2 = 1, && \text{(4.41a)} \\
&P_s\mathrm{tr}(\boldsymbol{F}_2\boldsymbol{X}) \le \bar{\gamma}_e\left(\sigma_r^2\mathrm{tr}(\overline{\boldsymbol{Y}}_2\boldsymbol{X}) + \tau q + \tau\sigma_e^2\right), && \text{(4.41b)} \\
&\mathrm{tr}(\overline{\boldsymbol{\Phi}}\boldsymbol{X}) \le \tau P_r, && \text{(4.41c)} \\
&\boldsymbol{X} \succeq \boldsymbol{0},\ \tau \ge 0. && \text{(4.41d)}
\end{aligned}
$$

Proposition 4.4.2. (P1′.1-sub2-SDR) *must yield a rank-one solution, i.e.,* $\boldsymbol{X}^* = \boldsymbol{w}\boldsymbol{w}^*$, *such that* $\boldsymbol{w}^* = \mu\boldsymbol{\nu}_{\max}(\boldsymbol{Z}^*)$, *and*

$$
\boldsymbol{Z}^* = P_s\boldsymbol{F}_1 - \lambda^*\sigma_r^2\overline{\boldsymbol{Y}}_1 - \alpha^*P_s\boldsymbol{F}_2 + \alpha^*\bar{\gamma}_e\sigma_r^2\overline{\boldsymbol{Y}}_2 - \beta_0^*\overline{\boldsymbol{\Phi}}, \tag{4.42}
$$

where $\boldsymbol{\nu}_{\max}(\boldsymbol{Z}^*)$ *represents the eigenvector corresponding to the largest eigenvalue of*

$\boldsymbol{Z}^*$, *and* $\mu = \sqrt{\frac{P_r}{\mathrm{tr}(\overline{\boldsymbol{\Phi}})\boldsymbol{\nu}_{\max}(\boldsymbol{Z}^*)\boldsymbol{\nu}_{\max}^H(\boldsymbol{Z}^*)}}$. *Also,* λ^*, α^* *and* β_0^* *are the optimal dual variables associated with* (4.41a)–(4.41c)*, respectively.*

Proof. See Appendix D. □

The only problem involved in Proposition 4.4.2 is the dual problem of (P1′.1-sub2-SDR), which admits a much simpler structure to solve than the primal one.

4.5 Joint AN-AF Beamforming with Imperfect CSI

4.5.1 Problem Formulation for Imperfect CSI

We use a deterministic spherical model [29, 49] to characterize the resulting CSIs' uncertainties such that

$$\mathcal{G}_0 = \{\boldsymbol{g}_0 | \boldsymbol{g}_0 = \hat{\boldsymbol{g}}_0 + \Delta\boldsymbol{g}_0, \Delta\boldsymbol{g}_0^H \boldsymbol{W}_0 \Delta\boldsymbol{g}_0 \leq 1\}, \tag{4.43a}$$

$$\mathcal{G}_k = \{\boldsymbol{g}_k | \boldsymbol{g}_k = \hat{\boldsymbol{g}}_k + \Delta\boldsymbol{g}_k, \Delta\boldsymbol{g}_k^H \boldsymbol{W}_k \Delta\boldsymbol{g}_k \leq 1\}, \forall k, \tag{4.43b}$$

$$\tilde{\mathcal{H}}_0 = \{\tilde{\boldsymbol{h}}_0 | \tilde{\boldsymbol{h}}_0 = \hat{\tilde{\boldsymbol{h}}}_0 + \Delta\tilde{\boldsymbol{h}}_0, \Delta\tilde{\boldsymbol{h}}_0^H \boldsymbol{W}_0' \Delta\tilde{\boldsymbol{h}}_0 \leq 1\}, \tag{4.43c}$$

$$\tilde{\mathcal{H}}_k = \{\tilde{\boldsymbol{h}}_k | \tilde{\boldsymbol{h}}_k = \hat{\tilde{\boldsymbol{h}}}_k + \Delta\tilde{\boldsymbol{h}}_k, \Delta\tilde{\boldsymbol{h}}_k^H \boldsymbol{W}_k'' \Delta\tilde{\boldsymbol{h}}_k \leq 1\}, \forall k, \tag{4.43d}$$

$$\mathcal{H}_k = \{\boldsymbol{h}_k | \boldsymbol{h}_k = \hat{\boldsymbol{h}}_k + \Delta\boldsymbol{h}_k, \Delta\boldsymbol{h}_k^H \boldsymbol{W}_k' \Delta\boldsymbol{h}_k \leq 1\}, \forall k, \tag{4.43e}$$

where $\hat{\boldsymbol{g}}_0$, $\hat{\boldsymbol{g}}_k$'s, $\hat{\tilde{\boldsymbol{h}}}_0$, $\hat{\tilde{\boldsymbol{h}}}_k$'s and $\hat{\boldsymbol{h}}_k$'s are the estimates of the corresponding channels; $\Delta\boldsymbol{g}_0$, $\Delta\boldsymbol{g}_k$'s, $\Delta\tilde{\boldsymbol{h}}_0$, $\Delta\tilde{\boldsymbol{h}}_k$'s and $\Delta\boldsymbol{h}_k$'s are their respective channel errors; the matrices $\boldsymbol{W}_0$, $\boldsymbol{W}_k$'s, $\boldsymbol{W}_0'$, $\boldsymbol{W}_k''$'s and $\boldsymbol{W}_k'$'s determine the shape of each error region. W.l.o.g., we set $\boldsymbol{W}_0 = \boldsymbol{I}/\epsilon_0$, $\boldsymbol{W}_0' = \boldsymbol{I}/\epsilon_0'$, $\boldsymbol{W}_k = \boldsymbol{I}/\epsilon_k$, $\boldsymbol{W}_k' = \boldsymbol{I}/\epsilon_k'$ and $\boldsymbol{W}_k'' = \boldsymbol{I}/\epsilon_k''$ for simplicity, where ϵ_0, ϵ_0', ϵ_k, ϵ_k', and ϵ_k'' represent the respective size of the bounded error regions, $k = 1, \ldots, K$.

Accordingly, we denote the robust counterpart for (P1′) as

$$
\begin{aligned}
(\mathrm{P2}'): \quad & \max_{\{\boldsymbol{Q}_k\},\boldsymbol{W}} \min_{\substack{\tilde{\boldsymbol{h}}_0\in\tilde{\mathcal{H}}_0,\tilde{\boldsymbol{h}}_k\in\tilde{\mathcal{H}}_k,\forall k \\ \boldsymbol{g}_0\in\boldsymbol{G}_0,\tilde{\boldsymbol{g}}_k\in\tilde{\mathcal{G}}_k,\forall k}} \bar{F}(\{\boldsymbol{Q}_k\},\boldsymbol{W}) \\
\mathtt{s.t.} \quad & \mathrm{tr}\left(\boldsymbol{W}\left(P_s\boldsymbol{h}_0\boldsymbol{h}_0^H+\sigma_r^2\boldsymbol{I}\right)\boldsymbol{W}^H\right)\le P_r, && (4.44a)\\
& \mathrm{tr}\left(\boldsymbol{Q}_k\right)\le \eta P_s \min_{\boldsymbol{h}_k\in\mathcal{H}_k,\forall k}\|\boldsymbol{h}_k\|^2,\ \forall k, && (4.44b)\\
& \boldsymbol{Q}_k\succeq\boldsymbol{0},\ \forall k. && (4.44c)
\end{aligned}
$$

An equivalent robust reformulation of (P1′.2) is given by

$$
(\mathrm{P2}'.2): \max_{\bar{\gamma}_e>0} \quad \frac{1+\hat{H}(\bar{\gamma}_e)}{1+\hat{F}(\bar{\gamma}_e)}, \tag{4.45}
$$

where $\hat{F}(\bar{\gamma}_e)=\bar{\gamma}_e$ and $\hat{H}(\bar{\gamma}_e)$ denotes the optimal value of problem (P2′.1) that is given by

$$
\begin{aligned}
&(\mathrm{P2}'.1):\\
&\max_{\boldsymbol{X},\{\boldsymbol{Q}_k\}} \min_{\substack{\tilde{\boldsymbol{h}}_k\in\tilde{\mathcal{H}}_k,\forall k \\ \tilde{\boldsymbol{h}}_0\in\tilde{\mathcal{H}}_0}} \frac{P_s\mathrm{tr}(\boldsymbol{F}_1\boldsymbol{X})}{\sigma_r^2\mathrm{tr}(\overline{\boldsymbol{Y}}_1\boldsymbol{X})+\sum_{k=1}^K\tilde{\boldsymbol{h}}_k^T\boldsymbol{Q}_k\tilde{\boldsymbol{h}}_k^{\dagger}+\sigma_b^2}\\
&\mathtt{s.t.} \max_{\substack{\boldsymbol{g}_k\in\mathcal{G}_k,\forall k \\ \boldsymbol{g}_0\in\mathcal{G}_0}} \frac{P_s\mathrm{tr}(\boldsymbol{F}_2\boldsymbol{X})}{\sigma_r^2\mathrm{tr}(\overline{\boldsymbol{Y}}_2\boldsymbol{X})+\sum_{k=1}^K\boldsymbol{g}_k^T\boldsymbol{Q}_k\boldsymbol{g}_k^{\dagger}+\sigma_e^2}\le\bar{\gamma}_e, && (4.46a)\\
&\mathrm{tr}(\overline{\boldsymbol{\Phi}}\boldsymbol{X})\le P_r, && (4.46b)\\
&\mathrm{tr}\left(\boldsymbol{Q}_k\right)\le\eta P_s\min_{\boldsymbol{h}_k\in\mathcal{H}_k,\forall k}\|\boldsymbol{h}_k\|^2,\ \forall k, && (4.46c)\\
&\mathrm{rank}(\boldsymbol{X})=1, && (4.46d)\\
&\boldsymbol{X}\succeq\boldsymbol{0},\ \boldsymbol{Q}_k\succeq\boldsymbol{0},\ \forall k. && (4.46e)
\end{aligned}
$$

As stated in Lemma 4.4.1, similarly, (P2′) can be proved to have the same optimal value as (P2′.2) and the same optimal solution as (P2′.1) when $\bar{\gamma}_e$ takes its optimal value to (P2′.2). As a result, (P2′) can be solved in a two-stage fashion as well. Specifically, given any $\bar{\gamma}_e$, we first solve (P2′.1) to obtain $\hat{H}(\bar{\gamma}_e)$ and then search for

the optimal $\bar{\gamma}_e$ to (P2′.2).

4.5.2 Solutions for Imperfect CSI

By ignoring (4.46d), (P2′.1) is recast as

$$\begin{aligned}&\text{(P2′.1-RW-SDR-Eqv)}:\\&\max_{\boldsymbol{X},\{\boldsymbol{Q}_k\}}\ \min_{\substack{\tilde{\boldsymbol{h}}_k\in\tilde{\mathcal{H}}_k,\forall k\\ \tilde{\boldsymbol{h}}_0\in\tilde{\mathcal{H}}_0}}\ \frac{P_s\mathrm{tr}(\boldsymbol{F}_1\boldsymbol{X})}{\sigma_r^2\mathrm{tr}(\overline{\boldsymbol{Y}}_1\boldsymbol{X})+\sum_{k=1}^K\tilde{\boldsymbol{h}}_k^T\boldsymbol{Q}_k\tilde{\boldsymbol{h}}_k^\dagger+\sigma_b^2}\\&\mathtt{s.t.}(4.46\text{a})-(4.46\text{c}),\ (4.46\text{e}).\end{aligned}\tag{4.47a}$$

It is worth noting that due to the rank-one relaxation of (P2′.1-RW-SDR-Eqv), the solution provided by (P2′.1-RW-SDR-Eqv) in general yields an upper-bound for $\hat{H}(\bar{\gamma}_e)$, which may not be achievable. However, in the sequel we insist on solving (P2′.1-RW-SDR-Eqv) that is regarded as an upper-bound benchmark for our proposed problem detailed later in this subsection.

Solutions to (P2′.1-RW-SDR-Eqv)

To make the "max-min" objective function of (4.47) tractable, we first rewrite (4.47) by the equivalent epigraph formulation as

$$\begin{aligned}&\text{(P2′.1-RW-SDR-Eqv)}:\\&\max_{\boldsymbol{X},\{\boldsymbol{Q}_k\},\delta}\ \delta\\&\mathtt{s.t.}\ \min_{\substack{\tilde{\boldsymbol{h}}_k\in\tilde{\mathcal{H}}_k,\forall k\\ \tilde{\boldsymbol{h}}_0\in\tilde{\mathcal{H}}_0}}\ \frac{P_s\mathrm{tr}(\boldsymbol{F}_1\boldsymbol{X})}{\sigma_r^2\mathrm{tr}(\overline{\boldsymbol{Y}}_1\boldsymbol{X})+\sum_{k=1}^K\tilde{\boldsymbol{h}}_k^T\boldsymbol{Q}_k\tilde{\boldsymbol{h}}_k^\dagger+\sigma_b^2}\geq\delta,\end{aligned}\tag{4.48a}$$

$$(4.46\text{a})-(4.46\text{c}),\ (4.46\text{e}).\tag{4.48b}$$

As there are potentially infinite number of constraints in (4.48a), (4.46a), and (4.46c), they are semi-indefinite and thus intractable. In the following, we equivalently

transform these constraints to tractable ones using *S-Procedure* and a generalized *S-Procedure* given in Lemmas 4.5.1 and 4.5.2, respectively.

Lemma 4.5.1. *(*S-Procedure *[84]) Let $f_m(\boldsymbol{x})$, $m = 1, 2$ be defined as*

$$f_m(\boldsymbol{x}) = \boldsymbol{x}^H \boldsymbol{A}_m \boldsymbol{x} + 2\Re\{\boldsymbol{b}_m^H \boldsymbol{x}\} + c_m, \tag{4.49}$$

where $\boldsymbol{A}_m = \boldsymbol{A}_m^H \in \mathbb{C}^{N\times N}$, $\boldsymbol{b}_m \in \mathbb{C}^{N\times 1}$ and $c_m \in \mathbb{R}$, and $\Re$ gives the real part of the input entity. Then the implication $f_1(\boldsymbol{x}) \geq 0 \Rightarrow f_2(\boldsymbol{x}) \geq 0$ holds if and only if there exists $\delta \geq 0$ such that

$$\begin{bmatrix} \boldsymbol{A}_2 & \boldsymbol{b}_2 \\ \boldsymbol{b}_2^H & c_2 \end{bmatrix} - \delta \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{b}_1 \\ \boldsymbol{b}_1^H & c_1 \end{bmatrix} \succeq \boldsymbol{0}, \tag{4.50}$$

provided there exists a point $\hat{\boldsymbol{x}}$ such that $f_m(\hat{\boldsymbol{x}}) > 0$, $m = 1, 2$.

Lemma 4.5.2. *([95,* Theorem 3.5*]) The robust block quadratic matrix inequality (QMI),*

$$\begin{bmatrix} \boldsymbol{H} & \boldsymbol{F} + \boldsymbol{G}\boldsymbol{X} \\ (\boldsymbol{F} + \boldsymbol{G}\boldsymbol{X})^H & \boldsymbol{C} + \boldsymbol{X}^H\boldsymbol{B} + \boldsymbol{B}^H\boldsymbol{X} + \boldsymbol{X}^H\boldsymbol{A}\boldsymbol{X} \end{bmatrix} \succeq \boldsymbol{0}, \quad \textit{for all} \boldsymbol{I} - \boldsymbol{X}^H\boldsymbol{D}\boldsymbol{X} \succeq \boldsymbol{0}, \tag{4.51}$$

is equivalent to

$$\exists t \geq 0, \textit{ such that } \begin{bmatrix} \boldsymbol{H} & \boldsymbol{F} & \boldsymbol{G} \\ \boldsymbol{F}^H & \boldsymbol{C} & \boldsymbol{B}^H \\ \boldsymbol{G}^H & \boldsymbol{B} & \boldsymbol{A} \end{bmatrix} - t \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{D} \end{bmatrix} \succeq \boldsymbol{0}. \tag{4.52}$$

First, by rearranging terms, (4.48a) can be equivalently transformed into the following linear form:

$$\min_{\substack{\tilde{\boldsymbol{h}}_k \in \tilde{\mathcal{H}}_k, \forall k \\ \tilde{\boldsymbol{h}}_0 \in \tilde{\mathcal{H}}_0}} P_s \mathrm{tr}(\boldsymbol{F}_1 \boldsymbol{X}) - \delta\sigma_r^2 \mathrm{tr}(\overline{\boldsymbol{Y}}_1 \boldsymbol{X}) - \delta \sum_{k=1}^{K} \tilde{\boldsymbol{h}}_k^T \boldsymbol{Q}_k \tilde{\boldsymbol{h}}_k^\dagger - \delta\sigma_b^2 \geq 0. \tag{4.53}$$

Recalling the following matrix equalities in line with our definition of vec$(\cdot)$ operation:

$$\mathrm{tr}(\boldsymbol{A}\boldsymbol{B}^T) = \mathrm{vec}^T(\boldsymbol{A})\mathrm{vec}(\boldsymbol{B}), \tag{4.54}$$

$$\mathrm{vec}(\boldsymbol{A}\boldsymbol{X}\boldsymbol{B}) = (\boldsymbol{A}\otimes\boldsymbol{B}^T)\mathrm{vec}(\boldsymbol{X}), \tag{4.55}$$

$$(\boldsymbol{A}\otimes\boldsymbol{B})^T = \boldsymbol{A}^T\otimes\boldsymbol{B}^T, \tag{4.56}$$

it follows that

$$\mathrm{tr}(\boldsymbol{F}_1\boldsymbol{X}) = \tilde{\boldsymbol{h}}^T(\boldsymbol{h}_0\otimes\boldsymbol{I})X(\boldsymbol{h}_0^H\otimes\boldsymbol{I})\tilde{\boldsymbol{h}}^\dagger, \tag{4.57}$$

$$\mathrm{tr}(\overline{\boldsymbol{Y}}_1\boldsymbol{X}) = \tilde{\boldsymbol{h}}^T(\boldsymbol{I}\otimes\boldsymbol{X})\tilde{\boldsymbol{h}}^\dagger, \tag{4.58}$$

where $\tilde{\boldsymbol{h}}\in\mathbb{C}^{N_t^3\times 1} = \mathrm{vec}(\tilde{\boldsymbol{h}}_0^T\otimes\boldsymbol{I})$. The equivalent channel model for $\tilde{\boldsymbol{h}}$ is given by $\tilde{\boldsymbol{h}} = \hat{\tilde{\boldsymbol{h}}}+\Delta\tilde{\boldsymbol{h}}$, where $\|\Delta\tilde{\boldsymbol{h}}\|^2\le N_t\epsilon_0'$ (c.f. (4.43)). By introducing $\boldsymbol{X}'' = (\boldsymbol{h}_0\otimes\boldsymbol{I})\boldsymbol{X}(\boldsymbol{h}_0^H\otimes\boldsymbol{I})$ and $\boldsymbol{X}' = \boldsymbol{I}\otimes\boldsymbol{X}$, (4.53) can thus be recast as

$$\min_{\substack{\tilde{\boldsymbol{h}}_k\in\tilde{\mathcal{H}}_k,\forall k\\ \tilde{\boldsymbol{h}}_0\in\tilde{\mathcal{H}}_0}} \Delta\tilde{\boldsymbol{h}}^T(P_s\boldsymbol{X}''-\delta\sigma_r^2\boldsymbol{X}')\Delta\tilde{\boldsymbol{h}}^\dagger + 2\Re\{\Delta\tilde{\boldsymbol{h}}^T(P_s\boldsymbol{X}''-\delta\sigma_r^2\boldsymbol{X}')\hat{\tilde{\boldsymbol{h}}}^\dagger\}$$
$$-\delta\sum_{k=1}^K\tilde{\boldsymbol{h}}_k^T\boldsymbol{Q}_k\tilde{\boldsymbol{h}}_k^\dagger-\delta\sigma_b^2\ge 0. \tag{4.59}$$

Hence, according to Lemma 4.5.1, the implication $\|\Delta\tilde{\boldsymbol{h}}\|^2\le N_t\epsilon_0'\Rightarrow$ (4.59) holds if and only if there exists $w^{(0)}\ge 0$ such that the following linear matrix inequality (LMI) holds:

$$\begin{bmatrix}\boldsymbol{H}_1 & \boldsymbol{F}_1\\ \boldsymbol{F}_1^H & c_1\end{bmatrix}\succeq\boldsymbol{0}, \tag{4.60}$$

where $\boldsymbol{H}_1 = P_s\boldsymbol{X}''-\delta\sigma_r^2\boldsymbol{X}'+w^{(0)}\boldsymbol{I}$, $\boldsymbol{F}_1 = (P_s\boldsymbol{X}''-\delta\sigma_r^2\boldsymbol{X}')\hat{\tilde{\boldsymbol{h}}}^\dagger$ and $c_1 = \hat{\tilde{\boldsymbol{h}}}^T(P_s\boldsymbol{X}''-\delta\sigma_r^2\boldsymbol{X}')\hat{\tilde{\boldsymbol{h}}}^\dagger-\delta\sum_{k=1}^K\tilde{\boldsymbol{h}}_k^T\boldsymbol{Q}_k\tilde{\boldsymbol{h}}_k^\dagger-\delta\sigma_b^2-w^{(0)}N_t\epsilon_0'$. Now, (4.48a) has been equivalently reformulated as (4.60). To further cope with channel uncertainties with regards to $\tilde{\boldsymbol{h}}_k$'s such that (4.60) holds for $\tilde{\boldsymbol{h}}_k\in\tilde{\mathcal{H}}_k$, $k=1,\ldots,K$, we need the

following proposition.

Proposition 4.5.1. *The semi-indefinite constraint of* (4.59) *can be equivalently recast as the following block matrix inequality:*

$$\begin{bmatrix} \boldsymbol{H}_1^{(K)} & \boldsymbol{F}_1^{(K)} & \boldsymbol{G}_1^{(K)} \\ \boldsymbol{F}_1^{(K)H} & c_1^{(K)} & \boldsymbol{B}_1^{(K)H} \\ \boldsymbol{G}_1^{(K)H} & \boldsymbol{B}_1^{(K)} & \boldsymbol{A}_1^{(K)} \end{bmatrix} - w^{(K)} \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \frac{-\boldsymbol{I}}{\epsilon_K''} \end{bmatrix} \succeq \boldsymbol{0}, \tag{4.61}$$

where $\boldsymbol{H}_1^{(K)}$, $\boldsymbol{F}_1^{(K)}$ *and* $c_1^{(K)}$ *are recursively given by*

$$\boldsymbol{H}_1^{(k)} = \begin{cases} \begin{bmatrix} \boldsymbol{A}_1^{(k-1)} + \frac{w^{(k-1)}}{\epsilon_{k-1}''}\boldsymbol{I} & \boldsymbol{G}_1^{(k-1)H} \\ \boldsymbol{G}_1^{(k-1)} & \boldsymbol{H}_1^{(k-1)} \end{bmatrix}, & k > 1; \\ P_s\boldsymbol{X}'' - \delta\sigma_r^2\boldsymbol{X}' + w^{(0)}\boldsymbol{I}, & k = 1, \end{cases} \tag{4.62}$$

$$\boldsymbol{F}_1^{(k)} = \begin{cases} \begin{bmatrix} \boldsymbol{B}_1^{(k-1)} \\ \boldsymbol{F}_1^{(k-1)} \end{bmatrix}, & k > 1; \\ (P_s\boldsymbol{X}'' - \delta\sigma_r^2\boldsymbol{X}')\hat{\tilde{\boldsymbol{h}}}^{\dagger}, & k = 1, \end{cases} \tag{4.63}$$

$c_1^{(k)} = \hat{\tilde{\boldsymbol{h}}}^T(P_s\boldsymbol{X}'' - \delta\sigma_r^2\boldsymbol{X}')\hat{\tilde{\boldsymbol{h}}}^{\dagger} - \delta\sum_{j=1}^{k}\hat{\tilde{\boldsymbol{h}}}_j^T\boldsymbol{Q}_j\hat{\tilde{\boldsymbol{h}}}_j^{\dagger} - \delta\sum_{i=k+1}^{K}\tilde{\boldsymbol{h}}_i^T\boldsymbol{Q}_i\tilde{\boldsymbol{h}}_i^{\dagger} - \delta\sigma_b^2 - w^{(0)}N_t\epsilon_0' - \sum_{l=1}^{k-1}w^{(l)}$, $k = 1, \ldots, K$. *In addition,* $\boldsymbol{G}_1^{(k)} \in \mathbb{C}^{(N_t^3+(k-1)N_t)\times N_t} = \boldsymbol{0}$, $\boldsymbol{B}_1^{(k)} = -\delta\boldsymbol{Q}_k\hat{\tilde{\boldsymbol{h}}}_k^{\dagger}$, $\boldsymbol{A}_1^{(k)} = -\delta\boldsymbol{Q}_k$, $k = 1, \ldots, K$, *and* $\{w^{(k)} \geq 0\}$ *denote pertinent auxiliary variables.*

Proof. See Appendix E. □

Next, (4.46a) is rewritten as

$$\max_{\substack{\boldsymbol{g}_k \in \mathcal{G}_k, \forall k \\ \boldsymbol{g}_0 \in \mathcal{G}_0}} \boldsymbol{g}^T\left(P_s\boldsymbol{X}'' - \bar{\gamma}_e\sigma_r^2\boldsymbol{X}'\right)\boldsymbol{g}^{\dagger} - \bar{\gamma}_e\sum_{k=1}^{K}\boldsymbol{g}_k^T\boldsymbol{Q}_k\boldsymbol{g}_k^{\dagger} - \bar{\gamma}_e\sigma_e^2 \leq 0, \tag{4.64}$$

where $\boldsymbol{g} \in \mathbb{C}^{N_t^2\times 1} = \mathrm{vec}(\boldsymbol{g}_0^T \otimes \boldsymbol{I})$ and the equivalent imperfect channel model is given by $\boldsymbol{g} = \hat{\boldsymbol{g}} + \Delta\boldsymbol{g}$ such that $\|\Delta\boldsymbol{g}\|^2 \leq N_t\epsilon_0$.

Proposition 4.5.2. *The semi-indefinite constraint of* (4.64) *is satisfied if and only if there exists* $v^{(k)} \geq 0$, $k = 1, \ldots, K$, *such that the following block matrix inequality holds:*

$$\begin{bmatrix} \boldsymbol{H}_2^{(K)} & \boldsymbol{F}_2^{(K)} & \boldsymbol{G}_2^{(K)} \\ \boldsymbol{F}_2^{(K)H} & c_2^{(K)} & \boldsymbol{B}_2^{(K)H} \\ \boldsymbol{G}_2^{(K)H} & \boldsymbol{B}_2^{(K)} & \boldsymbol{A}_2^{(K)} \end{bmatrix} - v^{(K)} \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \frac{-\boldsymbol{I}}{\epsilon_K} \end{bmatrix} \succeq \boldsymbol{0}, \tag{4.65}$$

where $\boldsymbol{H}_2^{(K)}$, $\boldsymbol{F}_2^{(K)}$ *and* $c_2^{(K)}$ *are recursively given by*

$$\boldsymbol{H}_2^{(k)} = \begin{cases} \begin{bmatrix} \boldsymbol{A}_2^{(k-1)} + \frac{v^{(k-1)}}{\epsilon_{k-1}} \boldsymbol{I} & \boldsymbol{G}_2^{(k-1)H} \\ \boldsymbol{G}_2^{(k-1)} & \boldsymbol{H}_2^{(k-1)} \end{bmatrix}, & k > 1; \\ -P_s \boldsymbol{X}'' + \bar{\gamma}_e \sigma_r^2 \boldsymbol{X}' + v^{(0)} \boldsymbol{I}, & k = 1, \end{cases} \tag{4.66}$$

$$\boldsymbol{F}_2^{(k)} = \begin{cases} \begin{bmatrix} \boldsymbol{B}_2^{(k-1)} \\ \boldsymbol{F}_2^{(k-1)} \end{bmatrix}, & k > 1; \\ (-P_s \boldsymbol{X}'' + \bar{\gamma}_e \sigma_r^2 \boldsymbol{X}') \hat{\boldsymbol{g}}^\dagger, & k = 1, \end{cases} \tag{4.67}$$

$$\begin{aligned} c_2^{(k)} =& \hat{\boldsymbol{g}}^T (-P_s \boldsymbol{X}'' + \bar{\gamma}_e \sigma_r^2 \boldsymbol{X}') \hat{\boldsymbol{g}}^\dagger + \bar{\gamma}_e \sum_{j=1}^{k} \hat{\boldsymbol{g}}_j^T \boldsymbol{Q}_j \hat{\boldsymbol{g}}_j^\dagger + \bar{\gamma}_e \sum_{i=k+1}^{K} \boldsymbol{g}_i^T \boldsymbol{Q}_i \boldsymbol{g}_i^\dagger \\ & + \bar{\gamma}_e \sigma_e^2 - v^{(0)} N_t \epsilon_0 - \sum_{l=1}^{k-1} v^{(l)}, \end{aligned} \tag{4.68}$$

$k = 1, \ldots, K$. *Also,* $\boldsymbol{G}_2^{(k)} = \boldsymbol{G}_1^{(k)}$, $\boldsymbol{B}_2^{(k)} = \bar{\gamma}_e \boldsymbol{Q}_k \hat{\boldsymbol{g}}_k^\dagger$, *and* $\boldsymbol{A}_2^{(k)} = \bar{\gamma}_e \boldsymbol{Q}_k$, $k = 1, \ldots, K$.

Proof. See Appendix F. □

Last, we rewrite (4.46c) to facilitate the robust optimization against the errors introduced by $\Delta \boldsymbol{h}_k$'s. By applying Lemma 4.5.1, (4.46c) holds if and only if there

exists $\mu_k \geq 0$, $k = 1, \ldots, K$, such that the following LMI constraint is met:

$$\begin{bmatrix} \eta P_s \boldsymbol{I} + \mu_k \boldsymbol{I} & \eta P_s \hat{\boldsymbol{h}}_k \\ \eta P_s \hat{\boldsymbol{h}}_k^H & \eta P_s \|\hat{\boldsymbol{h}}_k\|_2^2 - \mathrm{tr}(\boldsymbol{Q}_k) - \mu_k \epsilon_k' \end{bmatrix} \succeq \boldsymbol{0}, \ \forall k. \tag{4.69}$$

As such, (P2′.1-RW-SDR-Eqv) is now simplified as

$$\begin{aligned} \text{(P2′.1-RW-SDR-Eqv)}: \quad & \max_{\boldsymbol{X}, \{\boldsymbol{Q}_k\}, \delta} \ \delta \\ \texttt{s.t.} \quad & (4.61), (4.65), (4.69), (4.46\text{b}), (4.46\text{e}). \end{aligned}$$

Because of the non-convex term such as $\delta \boldsymbol{X}'$ in (4.61), problem (P2′.1-RW-SDR-Eqv) remains very hard to solve. We thus use the bisection method [84] w.r.t. δ to solve it. However, using bisection in addition to solving (P2′.2) by one-dimension search over $\bar{\gamma}_e$ may lead to very high complexity. As a result, we propose an alternative problem to approximate $\hat{H}(\bar{\gamma}_e)$.

Solutions to (P2′.1-RW-SDR)

We propose to approximate $\hat{H}(\bar{\gamma}_e)$ by the optimum value of the following problem.

$$\text{(P2′.1-RW-SDR)}: \quad \max_{\boldsymbol{X}, \{\boldsymbol{Q}_k\}, \tau} \min_{\tilde{\boldsymbol{h}}_0 \in \tilde{\mathcal{H}}_0} P_s \mathrm{tr}(\boldsymbol{F}_1 \boldsymbol{X}) \tag{4.70a}$$

$$\texttt{s.t.} \max_{\substack{\tilde{\boldsymbol{h}}_k \in \tilde{\mathcal{H}}_k, \forall k \\ \tilde{\boldsymbol{h}}_0 \in \tilde{\mathcal{H}}_0}} \sigma_r^2 \mathrm{tr}(\overline{\boldsymbol{Y}}_1 \boldsymbol{X}) + \sum_{k=1}^{K} \tilde{\boldsymbol{h}}_k^T \boldsymbol{Q}_k \tilde{\boldsymbol{h}}_k^\dagger + \tau \sigma_b^2 \leq 1, \tag{4.70b}$$

$$\max_{\substack{\boldsymbol{g}_k \in \mathcal{G}_k, \forall k \\ \boldsymbol{g}_0 \in \mathcal{G}_0}} \frac{P_s \mathrm{tr}(\boldsymbol{F}_2 \boldsymbol{X})}{\sigma_r^2 \mathrm{tr}(\overline{\boldsymbol{Y}}_2 \boldsymbol{X}) + \sum_{k=1}^{K} \boldsymbol{g}_k^T \boldsymbol{Q}_k \boldsymbol{g}_k^\dagger + \tau \sigma_e^2} \leq \bar{\gamma}_e, \tag{4.70c}$$

$$\mathrm{tr}(\overline{\boldsymbol{\Phi}} \boldsymbol{X}) \leq \tau P_r, \tag{4.70d}$$

$$\mathrm{tr}\left(\boldsymbol{Q}_k\right) \leq \tau \eta P_s \min_{\boldsymbol{h}_k \in \mathcal{H}_k, \forall k} \|\boldsymbol{h}_k\|^2, \ \forall k, \tag{4.70e}$$

$$\boldsymbol{X} \succeq \boldsymbol{0}, \ \boldsymbol{Q}_k \succeq \boldsymbol{0}, \ \forall k, \ \tau \geq 0. \tag{4.70f}$$

Remark 4.5.1. *It is worth noting that as the numerator and the denominator of the objective function in* (P2′.1) *are coupled by common uncertainty* $\tilde{\boldsymbol{h}}_0$*, Charnes-Cooper transformation, in general, cannot be applied to realize equivalent decoupling. As a result,* (P2′.1-RW-SDR) *yields a more conservative approximation for* $\hat{H}(\bar{\gamma}_e)$ *than* (P2′.1-RW-SDR-Eqv)*. However, considering that* (P2′.1-RW-SDR) *needs to be solved only once for given* $\bar{\gamma}_e$ *in contrast with* (P2′.1-RW-SDR-Eqv) *requring isection over* δ*, we exploit it in the sequel. The effectiveness of this approximation will be evaluated in Section 4.6.2.*

To proceed, we rewrite (P2′.1-RW-SDR) as

$$\begin{aligned}
&\text{(P2}'\text{.1-RW-SDR)}: \max_{\boldsymbol{X},\{\boldsymbol{Q}_k\},\delta,\tau} \quad \delta \\
&\texttt{s.t.} \quad \min_{\tilde{\boldsymbol{h}}_0\in\tilde{\mathcal{H}}_0} P_s\mathrm{tr}(\boldsymbol{F}_1\boldsymbol{X}) \geq \delta, && \text{(4.71a)} \\
&\quad \text{(4.70b)–(4.70f)}. && \text{(4.71b)}
\end{aligned}$$

First, by rewriting $\boldsymbol{F} = \boldsymbol{f}_1^\dagger \boldsymbol{f}_1^T$, where $\boldsymbol{f}_1 = \hat{\boldsymbol{f}}_1 + \Delta\boldsymbol{f}_1$, in line with Lemma 4.5.1, the implication $\|\Delta\boldsymbol{f}_1\|^2 \leq \|\boldsymbol{h}_0\|^2\epsilon_0' \Rightarrow$ (4.71a) holds if and only if there exists $s'^{(0)} \geq 0$ such that the following LMI constraint is satisfied:

$$\begin{bmatrix} P_s\boldsymbol{X} + s'^{(0)}\boldsymbol{I} & P_s\boldsymbol{X}\hat{\boldsymbol{f}}_1^\dagger \\ P_s\hat{\boldsymbol{f}}_1^T\boldsymbol{X} & P_s\hat{\boldsymbol{f}}_1^T\boldsymbol{X}\hat{\boldsymbol{f}}_1^\dagger - s'^{(0)}\epsilon_0'\|\boldsymbol{h}_0\|_2^2 - \delta \end{bmatrix} \succeq \boldsymbol{0}. \tag{4.72}$$

Next, as $\mathrm{tr}(\overline{\boldsymbol{Y}}_1\boldsymbol{X}) = \boldsymbol{y}_1^T\boldsymbol{X}'\boldsymbol{y}_1^\dagger$ (c.f. (4.58)), where $\boldsymbol{y}_1 = \mathrm{vec}(\tilde{\boldsymbol{h}}_0^T \otimes \boldsymbol{I})$, after some manipulation, (4.70b) holds if and only if there exists $s''^{(0)} \geq 0$ such that

$$\begin{bmatrix} s''^{(0)}\boldsymbol{I} - \sigma_r^2\boldsymbol{X}' & -\sigma_r^2\boldsymbol{X}'\hat{\boldsymbol{y}}_1^\dagger \\ -\sigma_r^2\hat{\boldsymbol{y}}_1^T\boldsymbol{X}' & c \end{bmatrix} \succeq \boldsymbol{0}, \tag{4.73}$$

where $c = -\sigma_r^2\hat{\boldsymbol{y}}_1^T\boldsymbol{X}'\hat{\boldsymbol{y}}_1^\dagger - \sum_{k=1}^K \tilde{\boldsymbol{h}}_k^T\boldsymbol{Q}_k\tilde{\boldsymbol{h}}_k^\dagger - \tau\sigma_b^2 + 1 - s''^{(0)}N_t\epsilon_0'$. Then (4.70b) can be

rewritten as

$$(4.73) \text{ for } \tilde{\boldsymbol{h}}_k \in \tilde{\mathcal{H}}_k,\ \forall k, \tag{4.74}$$

which is handled by the following proposition.

Proposition 4.5.3. *The semi-indefinite constraints in* (4.74) *can be replaced by the following LMI constraint:*

$$\begin{bmatrix} \bar{\boldsymbol{H}}^{(K)} & \bar{\boldsymbol{F}}^{(K)} & \bar{\boldsymbol{G}}^{(K)} \\ \bar{\boldsymbol{F}}^{(K)H} & \bar{C}^{(K)} & \bar{\boldsymbol{B}}^{(K)H} \\ \bar{\boldsymbol{G}}^{(K)H} & \bar{\boldsymbol{B}}^{(K)} & \bar{\boldsymbol{A}}^{(K)} \end{bmatrix} - s''^{(K)} \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \frac{-\boldsymbol{I}}{\epsilon''_K} \end{bmatrix} \succeq \boldsymbol{0}, \tag{4.75}$$

where $\bar{\boldsymbol{H}}^{(K)}$ *and* $\bar{\boldsymbol{F}}^{(K)}$ *are recursively given by*

$$\begin{cases} \bar{\boldsymbol{H}}^{(k)} = \begin{bmatrix} \bar{\boldsymbol{A}}^{(k-1)} + \frac{s''^{(k-1)}\boldsymbol{I}}{\epsilon''_{k-1}} & \bar{\boldsymbol{G}}^{(k-1)H} \\ \bar{\boldsymbol{G}}^{(k-1)} & \bar{\boldsymbol{H}}^{(k-1)} \end{bmatrix}, \bar{\boldsymbol{F}}^{(k)} = \begin{bmatrix} \bar{\boldsymbol{B}}^{(k-1)} \\ \bar{\boldsymbol{F}}^{(k-1)} \end{bmatrix} \\ k = 2, \ldots, K; \\ \bar{\boldsymbol{H}}^{(1)} = s''^{(0)}\boldsymbol{I} - \sigma_r^2 \boldsymbol{X}', \bar{\boldsymbol{F}}^{(1)} = -\sigma_r^2 \boldsymbol{X}' \hat{\boldsymbol{y}}_1^{\dagger}, \end{cases} \tag{4.76}$$

where $\bar{\boldsymbol{G}}^{(k)} = \boldsymbol{G}_1^{(k)}$, $\bar{\boldsymbol{B}}^{(k)} = -\boldsymbol{Q}_k \hat{\tilde{\boldsymbol{h}}}_k^{\dagger}$, $\bar{\boldsymbol{A}}^{(k)} = -\boldsymbol{Q}_k$, $\bar{c}^{(k)} = -\sigma_r^2 \hat{\boldsymbol{y}}_1^T \boldsymbol{X}' \hat{\boldsymbol{y}}_1^{\dagger} - \Sigma_{j=1}^{k} \hat{\tilde{\boldsymbol{h}}}_j^T \boldsymbol{Q}_j \hat{\tilde{\boldsymbol{h}}}_j^{\dagger} - \Sigma_{i=k+1}^{K} \tilde{\boldsymbol{h}}_i^T \boldsymbol{Q}_i \tilde{\boldsymbol{h}}_i^{\dagger} - \tau\sigma_b^2 + 1 - s''^{(0)} N_t \epsilon'_0 - \Sigma_{l=1}^{k-1} s''^{(l)}$, $k = 1, \ldots, K$, *and* $\{s''^{(k)} \geq 0\}$ *denote the auxiliary variables.*

Proof. We only sketch the proof herein since it is quite similar to that of Proposition 4.5.1. First, apply Lemma 4.5.2 to (4.73) given $\tilde{\boldsymbol{h}}_k$'s, $k = 2, \ldots, K$, fixed and obtain an initial LMI. Next, manipulate the resulting LMI according to the property of Schur-Complements to facilitate using Lemma 4.5.2. Then, repeat this procedure until all the semi-indefinite constraints w.r.t. $\tilde{\boldsymbol{h}}_k$'s have been incorporated into an equivalent LMI. □

Proposition 4.5.4. *The constraint in* (4.70c) *is guaranteed if and only if there exists*

$s^{(k)} \geq 0$, $k = 1, \ldots, K$, *such that the following LMI holds:*

$$\begin{bmatrix} \boldsymbol{H}^{(K)} & \boldsymbol{F}^{(K)} & \boldsymbol{G}^{(K)} \\ \boldsymbol{F}^{(K)H} & c'^{(K)} & \boldsymbol{B}'^{(K)H} \\ \boldsymbol{G}^{(K)H} & \boldsymbol{B}'^{(K)} & \boldsymbol{A}'^{(K)} \end{bmatrix} - s^{(K)} \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \frac{-\boldsymbol{I}}{\epsilon_K} \end{bmatrix} \succeq \boldsymbol{0}, \tag{4.77}$$

where $\boldsymbol{H}^{(k)}$ *and* $\boldsymbol{F}^{(k)}$ *are recursively given by*

$$\begin{cases} \boldsymbol{H}^{(k)} = \begin{bmatrix} \boldsymbol{A}'^{(k-1)} + \frac{s^{(k-1)}\boldsymbol{I}}{\epsilon_{k-1}} & \boldsymbol{G}^{(k-1)H} \\ \boldsymbol{G}^{(k-1)} & \boldsymbol{H}^{(k-1)} \end{bmatrix}, \boldsymbol{F}^{(k)} = \begin{bmatrix} \boldsymbol{B}'^{(k-1)} \\ \boldsymbol{F}^{(k-1)} \end{bmatrix} \\ k = 2, \ldots, K; \\ \boldsymbol{H}^{(1)} = -P_s \boldsymbol{X}'' + \bar{\gamma}_e \sigma_r^2 \boldsymbol{X}' + s^{(0)} \boldsymbol{I}, \\ \boldsymbol{F}^{(1)} = \left(-P_s \boldsymbol{X}'' + \bar{\gamma}_e \sigma_r^2 \boldsymbol{X}'\right) \hat{\boldsymbol{g}}^{\dagger}, \end{cases} \tag{4.78}$$

in which $\boldsymbol{G}^{(k)} = \boldsymbol{G}_1^{(k)}$, $\boldsymbol{B}'^{(k)} = \bar{\gamma}_e \boldsymbol{Q}_k \hat{\boldsymbol{g}}_k^{\dagger}$, $\boldsymbol{A}'^{(k)} = \bar{\gamma}_e \boldsymbol{Q}_k$, $c'^{(k)} = \hat{\boldsymbol{g}}^T(-P_s \boldsymbol{X}'' + \bar{\gamma}_e \sigma_r^2 \boldsymbol{X}')\hat{\boldsymbol{g}}^{\dagger} + \bar{\gamma}_e \sum_{j=1}^{k} \hat{\boldsymbol{g}}_j^T \boldsymbol{Q}_j \hat{\boldsymbol{g}}_j^{\dagger} + \bar{\gamma}_e \sum_{i=k+1}^{K} \boldsymbol{g}_i^T \boldsymbol{Q}_i \boldsymbol{g}_i^{\dagger} + \bar{\gamma}_e \tau \sigma_e^2 - s^{(0)} N_t \epsilon_0 - \sum_{l=1}^{k-1} s^{(l)}$, $k = 1, \ldots, K$, *and* $\{s^{(k)} \geq 0\}$ *denote the auxiliary variables.*

Proof. It is observed that (4.70c) differs from (4.46a) in the only respect that σ_e^2 is replaced by $\tau\sigma_e^2$. Hence the proof for Proposition 4.5.2 can be directly applied herein by substituting $\tau\sigma_e^2$ for σ_e^2. □

Last, by replacing "ηP_s" in (4.46c) with "$\tau\eta P_s$" in (4.70e), (4.70e) can be replaced by a similar LMI as (4.69), denoted by (4.71e′), in which the pertinent auxiliary variables are denoted by $\{\mu_k \geq 0\}$.

Consequently, the equivalent reformulation for problem (P2′.1-RW-SDR) can be

summarized as

$$\text{(P2}'\text{.1-RW-SDR)}: \max_{\substack{\boldsymbol{X},\{\boldsymbol{Q}_k\},\delta,\tau,\\ s^{(0)},s'^{(0)},s''^{(0)},\\ \{s^{(k)}\},\{s''^{(k)}\},\{\mu_k\}}} \delta$$

$$\text{s.t.} \quad (4.72), (4.75), (4.77), (4.71\text{e}'), (4.70\text{d}), (4.70\text{f}),$$

$$s^{(0)} \geq 0,\ s'^{(0)} \geq 0,\ s''^{(0)} \geq 0, \tag{4.79a}$$

$$s^{(k)} \geq 0,\ s''^{(k)} \geq 0,\ \mu_k \geq 0,\ \forall k. \tag{4.79b}$$

4.5.3 Proposed Rank-One Solutions for Imperfect CSI

(P2′.1-RW-SDR) is convex and can be solved efficiently by convex optimization tools such as CVX. Next, we derive the Lagrangian of (P2′.1-RW-SDR). Note that in the following expression, we only consider the uncertainties regarding $\tilde{\boldsymbol{h}}_0$, $\boldsymbol{h}_k$'s, $\tilde{\boldsymbol{h}}_k$'s, $\boldsymbol{g}_0$ and $\boldsymbol{g}_k$'s when $K = 1$ for the purpose of simplicity and the results can be easily extended to the case of $K > 1$. Denote the dual variables associated with (4.70d), (4.72), (4.75) and (4.77) by β_0, $\boldsymbol{W}$, $\boldsymbol{V}$ and $\boldsymbol{Y}$, respectively. Then the partial Lagrangian of (P2′.1-RW-SDR) w.r.t. $\boldsymbol{X}$ is

$$L(\overline{\boldsymbol{\Omega}}) = \mathrm{tr}(\bar{\boldsymbol{A}}\boldsymbol{X}), \tag{4.80}$$

where $\overline{\boldsymbol{\Omega}}$ is the set of all primal and dual variables, and

$$\begin{aligned}\overline{\boldsymbol{A}} &= P_s\boldsymbol{W}_{1,1} + 2P_s\Re\{\hat{\boldsymbol{f}}_1^\dagger\boldsymbol{W}_{12}^T\} + P_s w_{2,2}\hat{\boldsymbol{F}}_1 - \sigma_r^2\sum_{i=1}^{N_t}\left(\boldsymbol{V}_{1,1}^{(i,i)} + 2\Re\{\overline{\boldsymbol{V}}_{1,2}^{(i,i)}\} + \overline{\boldsymbol{V}}_{2,2}^{(i,i)}\right)\\ &- 2P_s\Re\{(\boldsymbol{h}_0^H\otimes\boldsymbol{I})\overline{\boldsymbol{Y}}_{2,1}(\boldsymbol{h}_0\otimes\boldsymbol{I})\} - P_s(\boldsymbol{h}_0^H\otimes\boldsymbol{I})\boldsymbol{Y}_{1,1}(\boldsymbol{h}_0\otimes\boldsymbol{I}) + 2\sigma_r^2\bar{\gamma}_e\sum_{i=1}^{N_t^2}\Re\{\overline{\boldsymbol{Y}}_{2,1}^{(i,i)}\}\\ &- P_s y_{2,2}(\boldsymbol{h}_0^H\otimes\boldsymbol{I})\hat{\boldsymbol{g}}^\dagger\hat{\boldsymbol{g}}^T(\boldsymbol{h}_0\otimes\boldsymbol{I}) + \sigma_r^2\bar{\gamma}_e\sum_{i=1}^{N_t^2}\overline{\boldsymbol{Y}}_{2,2}^{(i,i)}.\end{aligned} \tag{4.81}$$

In (4.81), $\hat{\boldsymbol{F}}_1 = \hat{\boldsymbol{f}}_1^\dagger\hat{\boldsymbol{f}}_1^T$; $\boldsymbol{W}_{i,j}$, $i,j = 1,2$, $\boldsymbol{V}_{i,j}$, $i,j = 1,\ldots,3$ and $\boldsymbol{Y}_{i,j}$, $i,j = 1,\ldots,3$ are the block submatrices of $\boldsymbol{W} \in \mathbb{C}^{(N_t^2+1)\times(N_t^2+1)}$, $\boldsymbol{V} \in \mathbb{C}^{(N_t^3+N_t+1)\times(N_t^3+N_t+1)}$ and

$\boldsymbol{Y} \in \mathbb{C}^{(N_t^3+2N_t+1)\times(N_t^3+2N_t+1)}$ with the same size as block submatrices in (4.72), (4.75) and (4.77), respectively. Moreover, in (4.81), $\overline{\boldsymbol{V}}_{1,2} = \hat{\boldsymbol{y}}_1^\dagger \boldsymbol{V}_{1,2}^T$, $\overline{\boldsymbol{V}}_{2,2} = v_{2,2}\hat{\boldsymbol{y}}_1^\dagger \hat{\boldsymbol{y}}_1^T$, $\overline{\boldsymbol{Y}}_{2,1} = \hat{\boldsymbol{g}}^\dagger \boldsymbol{y}_{1,2}^T$ and $\overline{\boldsymbol{Y}}_{2,2} = y_{2,2}\hat{\boldsymbol{g}}^\dagger \hat{\boldsymbol{g}}^T$. Furthermore, $\boldsymbol{V}_{1,1}^{(i,i)}$, $\overline{\boldsymbol{V}}_{1,2}^{(i,i)}$ and $\overline{\boldsymbol{V}}_{2,2}^{(i,i)}$ are the ith block diagonal submatrices of $\boldsymbol{V}_{1,1} \in \mathbb{C}^{N_t^3\times N_t^3}$, $\overline{\boldsymbol{V}}_{1,2} \in \mathbb{C}^{N_t^3\times N_t^3}$ and $\overline{\boldsymbol{V}}_{2,2} \in \mathbb{C}^{N_t^3\times N_t^3}$, respectively; $\overline{\boldsymbol{Y}}_{2,1}^{(i,i)}$ and $\overline{\boldsymbol{Y}}_{2,2}^{(i,i)}$ are the ith block diagonal submatrices of $\boldsymbol{Y}_{2,1} \in \mathbb{C}^{N_t^3\times N_t^3}$, and $\overline{\boldsymbol{Y}}_{2,2} \in \mathbb{C}^{N_t^3\times N_t^3}$, respectively.

Proposition 4.5.5. *1. The optimal $\boldsymbol{X}^*$ to* (P2′.1-RW-SDR) *is expressed as*

$$\boldsymbol{X}^* = \sum_{n=1}^{N_t^2-\bar{r}_c} \bar{a}_n \bar{\boldsymbol{\eta}}_n \bar{\boldsymbol{\eta}}_n^H + \bar{b}\bar{\boldsymbol{\xi}}\bar{\boldsymbol{\xi}}^H, \tag{4.82}$$

where $\bar{a}_n \geq 0$, $\forall n$, $\bar{b} > 0$, and $\bar{\boldsymbol{\xi}} \in \mathbb{C}^{N_t^2\times 1}$ is a unit-norm vector orthogonal to $\bar{\boldsymbol{\Xi}}$ (c.f. (4.29)*);*

2. According to (4.82)*, if* $\mathrm{rank}(\boldsymbol{X}^*) > 1$*, i.e., there exists at least one $\bar{a}_n > 0$, we reconstruct a solution to problem* (P2′.1-RW-SDR) *using*

$$\hat{\boldsymbol{X}}^* = \bar{b}\bar{\boldsymbol{\xi}}\bar{\boldsymbol{\xi}}^H, \tag{4.83}$$

$$\hat{\delta}^* = \delta^*, \tag{4.84}$$

$$\hat{\tau}^* = \tau^*, \tag{4.85}$$

while $\{\hat{\boldsymbol{Q}}_k^*\}$ are obtained by solving the following feasibility problem provided that $\hat{\boldsymbol{X}}^*$, $\hat{\delta}^*$, and $\hat{\tau}^*$ are given by (4.83)*,* (4.84) *and* (4.85)*, respectively:*

$$\begin{aligned}
&\text{(P2′.1-RW-SDR-sub)}: \quad \underset{\substack{\{\boldsymbol{Q}_k\},s''^{(0)},\\ \{s''^{(k)}\},\{\mu_k\}}}{\texttt{max}} \quad 0\\
&\texttt{s.t.} \quad (4.75) \text{ given } \hat{\boldsymbol{X}}^*, \hat{\tau}^*, \ (4.71\mathrm{e}') \text{ given } \hat{\tau}^*,\\
&\boldsymbol{Q}_k \succeq \boldsymbol{0},\ \mu_k \geq 0,\ \forall k,\\
&s''^{(0)} \geq 0,\ s''^{(k)} \geq 0,\ \forall k.
\end{aligned}$$

Proof. According to the KKT conditions of (P2′.1-RW-SDR), we have $\bar{\boldsymbol{A}}^* \boldsymbol{X}^* = \boldsymbol{0}$,

where $\bar{\boldsymbol{A}}^*$ is given by (4.81). Define $\bar{\boldsymbol{C}}^* = \bar{\boldsymbol{A}}^* - w_{2,2}^* P_s \hat{\boldsymbol{F}}_1$ with $\text{rank}(\bar{\boldsymbol{C}}^*)$ denoted by $\bar{r}_c$. Then take the similar procedure as **Case I** and **Case II** in Appendix C, it can be obtained that $\boldsymbol{X}^* = \sum_{n=1}^{N_t^2-\bar{r}_c} \bar{a}_n \bar{\boldsymbol{\eta}}_n \bar{\boldsymbol{\eta}}_n^H + \bar{b}\bar{\boldsymbol{\xi}}\bar{\boldsymbol{\xi}}^H$.

Next, we prove the second half of Proposition 4.5.5. According to (4.83),

$$P_s \text{tr}(\hat{\boldsymbol{F}}_1 \hat{\boldsymbol{X}}^*) = P_s \text{tr}(\hat{\boldsymbol{F}}_1 \boldsymbol{X}^*) \geq \min_{\tilde{\boldsymbol{h}}_0 \in \tilde{\mathcal{H}}_0} P_s \text{tr}(\boldsymbol{F}_1 \boldsymbol{X}) \geq \delta^*, \tag{4.86}$$

and thus (4.71a) holds true, which implies that the same optimal value as (P2′.1-RW-SDR), i.e., δ^*, is achievable. However, since the constraint in (4.70c) is ignored, the global optimal $\bar{\gamma}_e^*$ for (P2′.2) via solving (P2′.1-RW-SDR) is probably violated in (P2′.1-RW-SDR-sub). For example, $\frac{P_s \text{tr}(\boldsymbol{F}_2 \hat{\boldsymbol{X}}^*)}{\sigma_r^2 \text{tr}(\overline{\boldsymbol{Y}}_2 \hat{\boldsymbol{X}}^*) + \sum_{k=1}^K \boldsymbol{g}_k^T \hat{\boldsymbol{Q}}_k^* \boldsymbol{g}_k^\dagger + \hat{\tau}^* \sigma_e^2} = \bar{\gamma}_e^0 \geq \hat{F}(\bar{\gamma}_e^*)$, which results in the actual objective value for (P2′.2), $\frac{1+\hat{H}(\bar{\gamma}_e^*)}{1+\bar{\gamma}_e^0}$ smaller than $\frac{1+\hat{H}(\bar{\gamma}_e^*)}{1+\hat{F}(\bar{\gamma}_e^*)}$, and thus suboptimal for (P2′). □

The scheme that solves (P2′) is summarized in Table 4.2.

4.6 Numerical Results

Here we provide numerical examples to validate our results. We assume a typical scenario where the K helpers are evenly distributed around Alice with a radius of $\rho_k = 2\text{m}$ and $\theta_k = \frac{2\pi(k-1)}{K}$ (radian by default), where θ_k is the angle of direction (w.r.t. the Alice-relay link by default) of the kth helper, $k = 1, \ldots, K$. Alice, Bob and Eve are, w.l.o.g., assumed to have the same distance away from the AF relay with their angle of direction π, $\pi/6$ and $11\pi/6$, respectively. We also assume channel models with both large-scale fading, i.e., path loss and shadowing, and small-scale fading, i.e., multi-path fading. The simplified large-scale fading model is given by [96]

$$D = z A_0 \left(\frac{d}{d_0}\right)^{-\alpha}, \text{ for } d \geq d_0, \tag{4.87}$$

where z is a log-normal RV capturing the effect of shadowing with the standard derivation $\sigma = 4\text{dB}$, $A_0 = 10^{-3}$, d is the distance, d_0 is a reference distance set to be

Table 4.2: Algorithm II for (P2′)

- **Initialize** $\bar{\gamma}'_{e\_\text{search}} = 0 : \alpha' : \bar{\gamma}'_{e\,\max}$ and $i = 0$
- **Repeat**
  1) **Set** $i = i + 1$;
  2) Given $\bar{\gamma}_e = \bar{\gamma}'_{e\_\text{search}}(i)$,
     solve (P2′.1-RW-SDR) and **obtain** $\hat{H}(\bar{\gamma}_e^{(i)})$.
- **Until** $i = L'$, where $L' = \lfloor \frac{\bar{\gamma}'_{e\,\max}}{\alpha'} \rfloor + 1$ is the length of $\bar{\gamma}'_{e\_\text{search}}$
- **Set** $\bar{\gamma}_e^* = \bar{\gamma}'_{e\_\text{search}} \left(\arg\max_i \left\{ \frac{1+\hat{H}(\bar{\gamma}_e^{(i)})}{1+\bar{\gamma}_e^{(i)}} \right\} \right)$ for (P2′.2)
- Given $\bar{\gamma}_e^*$, **solve** (P2′.1-RW-SDR) to obtain $(\boldsymbol{X}^*, \{\boldsymbol{Q}_k^*\}, \delta^*, \tau^*)$
  if $\text{rank}(\boldsymbol{X}^*) = 1$, **apply** EVD on $\boldsymbol{X}^*$ such that $\boldsymbol{X}^* = \boldsymbol{w}^* \boldsymbol{w}^{*H}$;
  else
  - **construct** $(\hat{\boldsymbol{X}}^*, \hat{\delta}^*, \hat{\tau}^*)$, according to (4.83)-(4.85) and **set** $\boldsymbol{w}^* = \sqrt{\bar{b}}\bar{\boldsymbol{\xi}}$;
  - given $\hat{\boldsymbol{X}}^*$, $\hat{\delta}^*$ and $\hat{\tau}^*$,
    obtain $\{\hat{\boldsymbol{Q}}_k^*\}$ by solving (P2′.1-RW-SDR-sub).

  end
- **Recover** $\boldsymbol{W}^* = \text{vec}^{-1}(\boldsymbol{w}^*)$

1m, and $\alpha = 2$ is the path loss exponent. Specifically, the channels including $\boldsymbol{h}_k$'s, $\boldsymbol{h}_0$, $\tilde{\boldsymbol{h}}_0$ and $\boldsymbol{g}_0$, are assumed to suffer from Rician fading while the channels from the HJ helpers to Bob ($\tilde{\boldsymbol{h}}_k$'s) and Eve ($\boldsymbol{g}_k$'s) follow Rayleigh distribution due to the missing of line-of-sight (LOS) components with their respective average gain specified by (4.87). Take $\boldsymbol{h}_k$, $\forall k$, as an example, $\boldsymbol{h}_k = \sqrt{\frac{K_R}{K_R+1}}\bar{\boldsymbol{h}}_k + \sqrt{\frac{1}{K_R+1}}\check{\boldsymbol{h}}_k$, where $\bar{\boldsymbol{h}}_k$ is the LOS component with $\|\bar{\boldsymbol{h}}_k\|_2^2 = D$ (c.f. (4.87)), $\check{\boldsymbol{h}}_k$ is the Rayleigh fading component denoted by $\check{\boldsymbol{h}}_k \sim \mathcal{CN}(0, D\boldsymbol{I})$, and K_R is the Rician factor set to be 3. Note that for the involved LOS component, we use the far-field uniform linear antenna array to model the channels [97]. In addition, unless otherwise specified, the number of HJ helpers, K is set to be 5; the AF relay is assumed to be 5m away from Alice; the EH efficiency, $\eta = 0.5$ and $\sigma_r^2 = \sigma_b^2 = \sigma_e^2 = -50$dBm. The results presented in Section 4.6.1 are obtained by averaging over 500 times of independent trials.

4.6.1 The Perfect CSI Case

We compare the proposed optimal solutions with three suboptimal schemes in the case of perfect CSI. One suboptimal scheme, denoted by "Suboptimal 1", is introduced in Section 4.4.3 by exploiting the optimal structure of $\boldsymbol{W}$. The other described in Section 4.4.3 is known as optimal null-space ZF, denoted by "Suboptimal 2". Specifically, each jamming beam $\boldsymbol{n}_k$ is restricted to lie in the orthogonal space of $\tilde{\boldsymbol{h}}_k^\dagger$ such that $\boldsymbol{n}_k$'s cause no interference to the IR while maximizing its effect of jamming at the eavesdropper. As a benchmark, we also present the well-known *isotropic jamming* that is particularly useful when there is no Eve's CSI known at each *HJ* helper, H_k, $\forall k$ [50], denoted by "Suboptimal 3". Note that the difference between "Suboptimal 2" and "Suboptimal 3" only lies in the design of jamming noise, for which the former also aligns the jamming noise to an equivalent Eve's channel to confront Eve with most interference, while the latter transmits isotropic jamming with $\tilde{\boldsymbol{n}}_k \sim \mathcal{CN}(\boldsymbol{0}, \eta P_s \|\boldsymbol{h}_k\|^2/(N_t-1))$, $k = 1, \ldots, K$, in directions orthogonal to $\tilde{\boldsymbol{h}}_k$'s, due to lack of knowledge of Eve's channel and thus is expected to be less efficient than "Suboptimal 2" with perfect CSI.

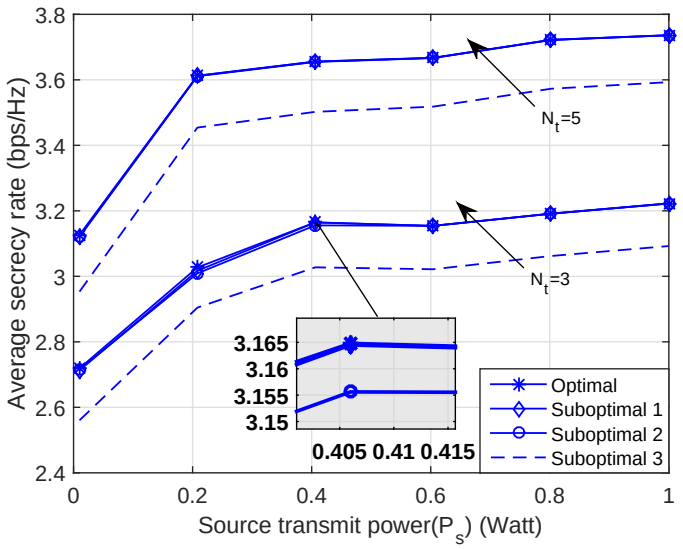


Figure 4.2: Secrecy rate versus Alice's transmit power with perfect CSI.

First, we study the secrecy rate at the receiver versus the transmit power of the transmitter, P_s with $P_r = 10$dBm. Fig. 4.2 demonstrates that for both cases of $N_t = 3$ and $N_t = 5$, the average secrecy rate increases and tends to be saturated as P_s goes to 30dBm. It also illustrates that "suboptimal 1" and "suboptimal 2" closely approach the optimal solutions while "Suboptimal 3" is outperformed more succinctly with larger number of antennas at the AF relay and the HJ helpers. Moreover, with N_t increasing, the average secrecy rate gets larger as a result of the higher array gain of the AF relay and more available power transferred to the HJ helpers.

In addition, we show in Fig. 4.3 the secrecy rate achieved by different schemes versus the transmit power of the AF relay, P_r with $P_s = 30$dBm. It is seen that the average secrecy rate first grows faster and then slower, since when P_r increases, not only the desired signal but also the noise yielded from the first transmission phase is amplified to a larger extent. In addition, the performance gap between the optimal scheme and suboptimal schemes is almost negligible. Similar to Fig. 4.2, "Suboptimal 3" appears to have certain performance loss from the optimality but is considered as a promising scheme when no Eve's CSI is available at the HJ helpers.

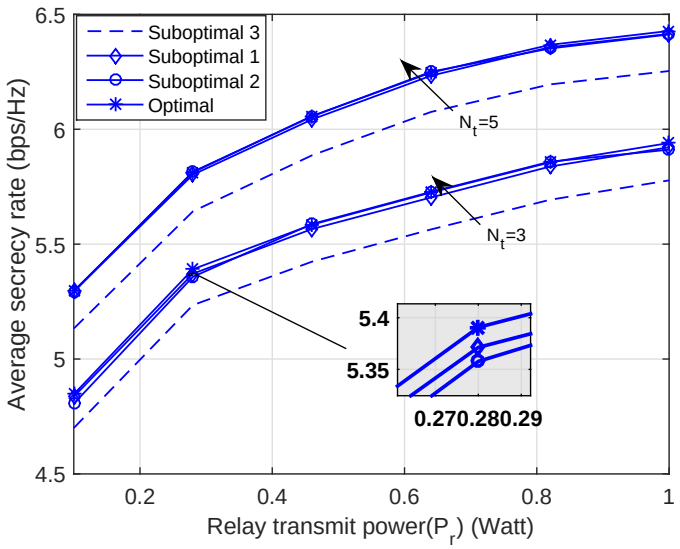


Figure 4.3: Secrecy rate versus the relay's transmit power with perfect CSI.

4.6.2 The Imperfect CSI Case

Now, we consider the imperfect CSI case and compare the proposed scheme *Robust SDR with HJ*, which is obtained by solving (P2′.1-RW-SDR-sub), against some benchmarks. Note that there are two upper-bound benchmark schemes, namely, *Robust SDR with HJ* and *Robust-eqv with HJ*, as well as two lower-bound benchmarks, which are *Robust w/o HJ* and *Non-robust with HJ*. For *Robust SDR with HJ* (*Robust-eqv with HJ*), given any $\bar{\gamma}_e$, $\hat{H}(\bar{\gamma}_e)$ is approximated by solving the rank constraint relaxed problem (P2′.1-RW-SDR) ((P2′.1-RW-SDR-Eqv)). On the other hand, for *Robust w/o HJ*, we solve (P2′.1-RW-SDR) by setting $\boldsymbol{Q}_k = 0$, $\forall k$ while for *Non-robust with HJ*, (4.11) is evaluated by applying the optimal solutions to (P1′.1) assuming perfect CSI, to the actual channels including errors that are generated from the sets defined in (4.43).

To assess the worst-case secrecy performance, we use the metric, namely, *secrecy outage probability*, defined as [39]:

$$p = P_r(r \leq r_0^*), \tag{4.88}$$

where r_0^* obtained by solving (P2′) is termed as the $100p\%$-*secrecy outage rate*.

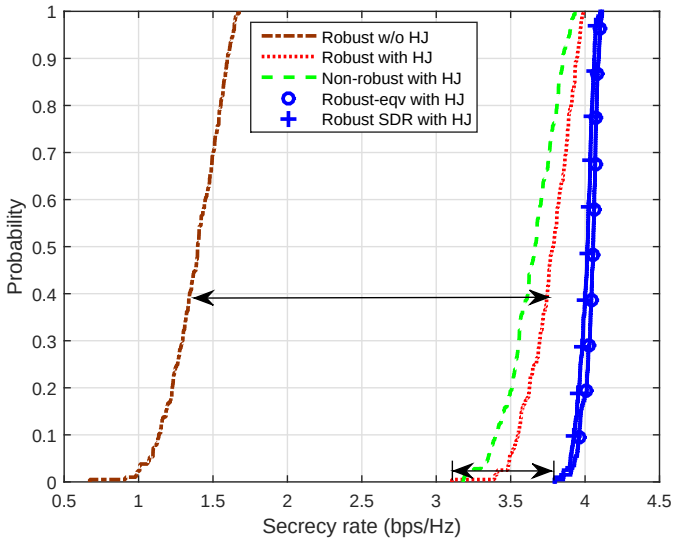


Figure 4.4: CDFs of the achievable secrecy rate.

The parameters are set identical to those in the perfect CSI case. Regarding the uncertainty model in (4.43), we introduce the uncertainty ratios associated with ϵ_0, ϵ_0', ϵ_k, ϵ_k' and ϵ_k'' as α_0, α_0', α_k, α_k' and α_k'', respectively. For instance, α_0 is

$$\alpha_0^2 = \frac{\epsilon_0}{\mathbb{E}[\|\boldsymbol{g}_0\|^2]}, \tag{4.89}$$

while α_0', α_k's, α_k''s and α_k'''s are similarly defined and thus omitted here for brevity. Besides, it is reasonable to assume that the channel estimates w.r.t Eve suffer from more errors than those for Alice and Bob. Hence, we set $\alpha_0'^2 = \alpha_k'^2 = \alpha_k''^2 = 1\%$ while $\alpha_0^2 = \alpha_k^2 = 10\%$, $k = 1, \ldots, K$, unless otherwise specified.

Fig. 4.4 demonstrates the cumulative density function (CDF) of the achievable secrecy rate from 1000 samples of random channel errors uniformly distributed over the sets defined by (4.43) given fixed actual channel realization. We set $P_r = 20$dBm, $P_s = 30$dBm, $N_t = 3$, $K = 5$ and $\alpha_0'^2 = \alpha_k'^2 = \alpha_k''^2 = 2\%$, $k = 1, \ldots, K$. Despite being suboptimal to the upper-bound schemes of "Robust SDR with HJ" and "Robust-eqv with HJ", the proposed "Robust with HJ" scheme outperforms its non-robust counterpart "Non-robust with HJ" particularly in the low range of probability, and overwhelmingly surpasses the "Robust w/o HJ". For example,

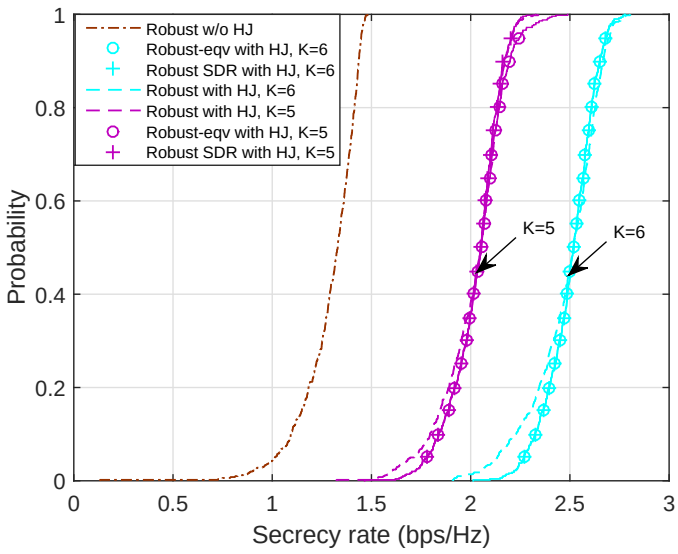


Figure 4.5: Secrecy outage probability for $K = 3$ and $K = 5$ HJ helpers, respectively.

"Robust with HJ" can achieve a secrecy rate of around 3.5bps/Hz in the 3% worst case versus that of 3.3bps/Hz and 1.0bps/Hz for the "Non-robust with HJ" and "Robust w/o HJ", respectively. The solutions for "Robust SDR with HJ" is also seen to admit very little gap from those for "Robust-eqv with HJ", which suggests that approximating $\hat{H}(\bar{\gamma}_e)$ by solving the complexity reduced "Robust SDR with HJ " leads almost no performance loss.

Fig. 4.5 illustrates the CDF of the achievable secrecy rate from 1000 samples of random channel errors generated in the same way as Fig. 4.4, with $P_r = 20$dBm, $P_s = 30$dBm and $N_t = 3$. It is observed that proposed solutions to "Robust with HJ" nearly achieve their upper-bound rank constraint relaxed solutions, i.e., SDR, to "Robust upper SDR with HJ" throughout the whole range of outage probability. Moreover, the "Robust w/o HJ" yields the worst performance. In particular, when the outage probability falls to 3%, the "Robust w/o HJ" achieves a worst-case secrecy rate of less than 1bps/Hz while the proposed scheme can still guarantee an outage rate of rough 1.64bps/Hz and 2.07bps/Hz for $K = 5$ and $K = 6$, respectively. Also, it is observed that increasing the number of HJ helpers will improve the secrecy performance, but we do not draw conclusions on the extent to which the secrecy rate can increase, since it also depends on the level of channel estimation inaccuracy. For

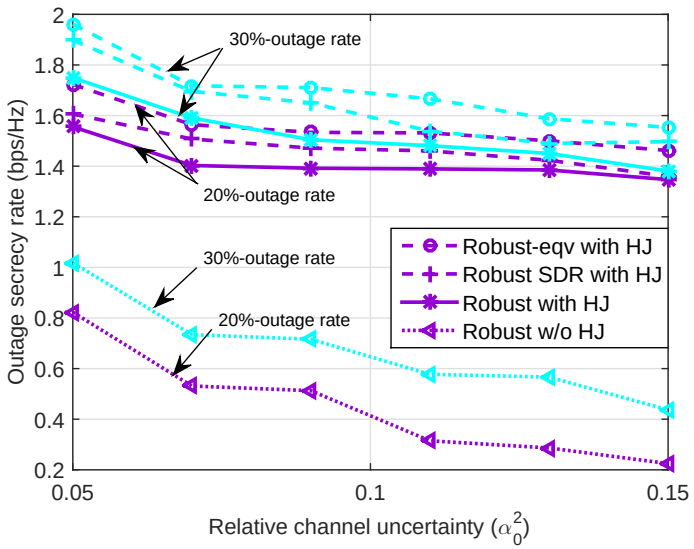


Figure 4.6: Secrecy outage rate versus the normalized channel errors.

example, more HJ helpers may also yield larger interference to the legitimate receiver if the channels from HJ helpers to Bob are not as well estimated as this instance of $\alpha_k''^2 = 1\%$, $\forall k$. Hence we suggest that in practice, a mild number of HJ helpers are sufficient in view of the trade-off between complexity and performance.

Fig. 4.6 shows two different levels ($p = 0.20$ and $p = 0.30$) of secrecy outage rate versus the channel uncertainty ratios (assuming $\alpha_0 = \alpha_k$, $k = 1, \ldots, K$), in which $P_r = 30$dBm, $P_s = 30$dBm, $N_t = 3$ and $K = 5$. It is observed that the secrecy outage rate by the proposed schemes decreases slowly with the eavesdropper's CSI error ratios, which validates the motivation of the worst-case robust optimization. It is worth noting that the advantage of the *HJ* protocol is more significant when the normalized channel uncertainty of Eve's channels surpasses 10%, since the *HJ* scheme provides more degree of freedom for robust design and thus capable of guaranteeing larger worst-case secrecy rate against worse channel conditions compared to that without *HJ*. The reasonably suboptimal performance of the proposed "Robust with HJ" is also seen as from Figs. 4.4 and 4.5.

Fig. 4.7 studies the $100p\%$-secrecy outage rate for $p = 0.05$ and $p = 0.20$, respectively, versus the transmit power of the AF relay. Specifically, we set $P_s = 30$dBm, $N_t = 3$, and $K = 5$. As observed similarly from Fig. 4.6, the robust

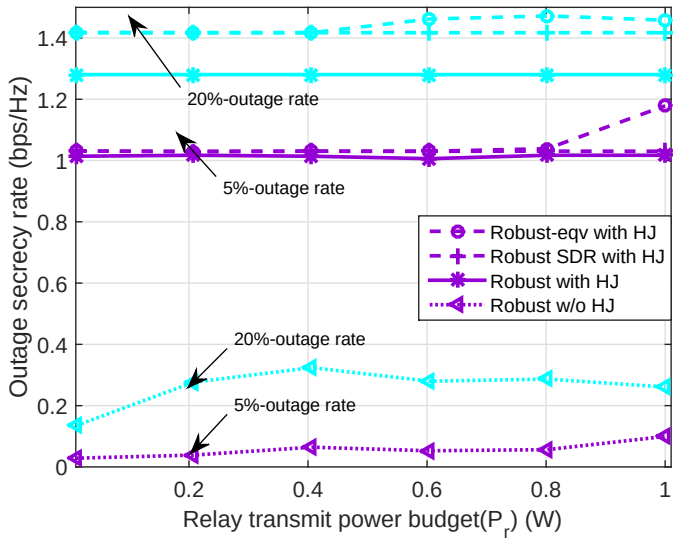


Figure 4.7: Secrecy outage rate versus the relay's transmit power.

schemes with the assistance of HJ helpers perform considerably better than solutions without HJ helpers. Furthermore, when the transmit power is set relatively large, i.e., $P_s = 30$dBm, it is seen that continuously increasing P_r does not contribute much to the secrecy performance, because in this situation the increased amplified noise at the AF relay compromises the performance, which provides useful insight for practical setting of P_r. In addition, the proposed "Robust with HJ" is observed striking a good trade-off between optimality and complexity compared with the two upper-bound solutions.

4.7 Chapter Summary

A novel *harvest-and-jam (HJ)* relaying protocol was considered in this chapter to improve the secret wireless communications in a multi-antenna AF relay wiretap channel. The AN covariance matrices at HJ helpers and the AF relay beamforming matrix have been jointly optimized to maximize the achievable secrecy rate and/or worst-case secrecy rate for the legitimate receiver subject to the transmit power constraints of the AF relay as well as the HJ helpers, under perfect and imperfect CSI assumptions, respectively, using the technique of SDR. The SDR were shown

tight for the perfect CSI case while a suboptimal rank-one reconstruction algorithm for the robust formulation under imperfect CSI was presented achieving promising trade-offs between complexity and performance. The effectiveness of the proposed schemes were also verified by numerical results.

Chapter 5

CJ-aided Multi-AF Relaying for Secrecy in SWIPT Networks

5.1 Introduction

Alike in Chapter 4, the benefits brought to PLS by WPCN in cooperative communications continues to be exploited in this chapter. It has been well investigated in [18, 22] that the secrecy capacity for MISO wiretap channel in the presence of one eavesdropper is achieved without necessity of employing AN, while that in the presence of multiple eavesdroppers can normally be increased by exploiting AN [33]. Analogous to MISO wiretap channel with multiple eavesdroppers, the multi-AF relaying with security concern in the second hop of a multi-hop cooperative network thus forms a *virtual* MISO wiretap channel, in which each single-antenna relay acts as an MISO antenna. As a result, AN, also known as CJ synthesized by external helpers in cooperative scenarios, will also contribute to increasing the secrecy capacity in the presence of multiple eavesdroppers [51]. Further, instead of implementing CJ with separate relays from those used for CB, a more general cooperation strategy is proposed that a source transmits confidential information to a legitimate Rx aided by a set of WEH-enabled AF relays operating with *CB mixed with CJ*. Specifically, each relay employs a heterogeneous PS protocol, which devices the PS receiver for SWIPT in the first transmission phase and further splits its harvested power for forwarding the received information and generating AN in the second transmission phase. The objective is thus to maximize the secrecy rate subject to self-sustained power constraints at the relays, by jointly optimizing the CB

and CJ using SDR techniques, which is proved to be tight. *Centralized* closed-form expressions for the relay beamforming have been derived for two practical types of WEH-enabled relays, *SPS* and *DPS*, respectively. In addition, a completely *distributed* algorithm assuming only local CSI available at each relay is provided as performance benchmark. Simulation results demonstrate the effectiveness of the proposed multi-AF CB mixed with CJ compared with other suboptimal designs. Furthermore, motivated by the advantage of large scale antennas, a large number of WEH-enabled single-antenna AF relays are exploited to simulate the effect of massive MISO, and the average secrecy rate based on the *matched filter (MF)* relay weights for fixed PS ratios is also analyzed utilizing law of large numbers (LLN).

The rest of the chapter is organized as follows. In Section 5.3, two types of WEH-enabled receiver architecture for the AF relays are described and the secrecy rate region of the relay wiretap channel is defined. The secrecy rate maximization problems that jointly optimize the AN and AF relay beamforming for WEH-enabled relays operating with the two types of receivers are respectively formulated in Section 5.4. The problems are respectively solved by centralized schemes in Section 5.5 and completely distributed approaches in Section 5.6. The proposed schemes are then verified by extensive simulations in Section 5.7. In addition, secrecy performance analysis for the well known MF relay beamforming when the number of relays goes to infinity is provided in Section 5.8 based on asymptotic results. Finally, the chapter is concluded in Section 5.9.

5.2 Related Work

5.2.1 Cooperation for PLS Enhancements

Among a variety of emerging application scenarios such as relay networks and device-to-device (D2D) communications etc., PHY-layer security enhancements by means of cooperative communications have begun drawing much attention [8, 21, 35, 36, 45–47, 50–53, 55, 56] since the relay-assisted secure transmission strategy was first

considered in [7, 43]. Related cooperative schemes including multiple relays against eavesdropping can be primarily classified into three categories [8]: *decode-and-forward (DF)*, *amplify-and-forward (AF)*, and *cooperative jamming (CJ)*. The robust secure beamforming designs were studied in [48] for a single multi-antenna AF relay-aided wiretap channel. Exploiting node cooperation in a DF fashion to achieve PHY-layer security was considered in [35, 47] by deriving optimal power allocations among nodes and/or multi-carrier resources. For CJ, coordinated CJ refers to the scheme of generating a common jamming signal across all single-antenna relay helpers [8, 21, 35, 36], while uncoordinated CJ assumes that each relay helper emits independent artificial noise (AN) to confound the eavesdroppers [50]. On the other hand, under the circumstances that direct link is broken between the Tx and the legitimate Rx, i.e., some of the relays have to take on their conventional role of forwarding the information while other spare ones can be employed for CJ, a selection of their function was performed in [52, 53]. Furthermore, a recent paradigm that generalizes all the above mentioned cooperation strategies is CB mixed with CJ [51, 55], into which this work also falls. This scheme splits the available power at each relay into two parts: one for forwarding the confidential message and the other for CJ.

5.2.2 WEH-enabled CB Mixed with CJ

In spite of taking full advantage of all relays' d.o.f, CB mixed with CJ may be prohibitive in applications with low power devices, mainly because these idle relays with limited battery supplies, prefer to saving power for their own traffic to assisting in other's secrecy transmission. Thanks to WPCN, SWIPT, nevertheless, provides essential incentives for these potential helpers to collaborate with others. However, since the received signal used for harvesting RF energy cannot be reused for decoding the modulated information due to hardware limitations [98], practical Rx architectures need to be designed to resolve this issue (see [9] for more detail). Typically, except for integrated receiver (IntRx) that splits the signal after converting it to DC current, which thus requires non-coherent detection, time

switching (TS)/power splitting (PS) receivers are more often implemented in terms of high-data-rate transmission that have their the power splitting unit installed in the RF front-ends of separate EH and information decoding (ID) receivers. TS receiver switches its operations periodically between EH and ID in the two time slots of one transmission block; RS receiver enables simultaneous wireless information transfer (WIT) and WPT by splitting the received power into two streams, where α portion of the received power in one stream is used for EH while the remaining $(1-\alpha)$ in the other stream is used for ID. Besides, antenna switching (AS) receiver is a special case of PS receiver in multi-antenna applications that simply connects each receiving antenna to either EH or ID receiver, the rate-energy region achieved by which has little gap from that of PS receiver when the number of antennas are large enough [99].

The minimum power optimization that incorporates the joint transmit beamforming and receiver PS ratios was considered in [100] for multiuser MISO interference channel, where the received power is not used for another hop of transmission. Facilitated by SWIPT and the d.o.f of multi-hop communications, cooperative relay(s) using the aforementioned Rx architectures were early investigated in [101] and then widely studied in [101–103] and [58, 73, 104], w/o and with secrecy transmission taken into account, respectively.The receiver architecture that is employed for the multi-AF relays in this chapter is DPS, which turns out to be the most general Rx operation that was initially proposed in [9].

Note that although the setting of the problem seems a special case of that considered in [51], its optimal CB mixed with CJ design is not applicable herein due to the multiplicative nature in beamforming weights incurred by adjustable PS ratios that intrinsically poses more intractability. Furthermore, in spite that an efficient algorithm that jointly optimizes PS ratios and AF relay beamforming was investigated in [73] to maximize the secrecy rate, this chapter differs from it in two folds. The most general CB mixed with CJ strategy is considered while [73] did not take the relay-based AN into account in its second transmission phase; the solutions derived

for the CB without CJ is proved to be global optimal whereas the algorithm proposed in [73] only converged to a local optimal solution.

5.2.3 PLS in a Large Scale

Massive MIMO or large scale MIMO has been considered to pave one of the main paths for cellular networks in 5G communications, in which base stations (BSs) are equipped with a magnitude of (more than) hundreds of antennas and thus provide multiplexing or diversity gain in a larger scale [105]. More strikingly, it has been shown that in such system serving multi-users in the same time-frequency resource, simple linear precoding schemes such as matched-filter (MF) and ZF actually achieve near performance to the capacity-achieving dirty-paper coding (DPC) [106] since the effect of noise and multiuser-interference are substantially reduced as the number of antennas grows to infinity. One potential advantage of massive MIMO for PLS is its considerable secrecy rate against the external eavesdropping since the received signal power at the intended user is several orders of magnitude higher than that at the external eavesdroppers. Zhu et al. in [107] studied the secrecy performance of MF precoding along with AN generation for multi-cell massive MIMO downlink transmission in the presence of an external multi-antenna eavesdropper. The authors investigated the conditions that guarantee secure transmission, among which, AN is in general required to achieve a positive ergodic secrecy rate. In addition, unlike that in multiuser-MIMO downlink transmission with finite number of antennas at the BS, random AN shaping matrices are promising alternative to *null-space* AN under the circumstances of no eavesdropper's CSIT in view of their perfrormance/complexity trade-off. Furthermore, compared with multiuser massive MIMO systems w/o secrecy consideration, the ergodic secrecy rate no longer monotonically increases with the number of BS antennas to the practical interest of pilot contamination [105].

On another front, exploiting cooperative strategy with massive MIMO relay has also aroused lots of research interest very recently. In [108], H. Q. Ngo et al. explored a massive-antenna-equipped DF relay operating in full-duplex mode over which

multiple sources communicate with their respective destinations. They showed that the advantage of full-duplex is especially prominent in the massive antenna array case thanks to the significantly reduced loop interference and asymptotically disappearing inter-pair interference plus noise. Moreover, secure cooperative transmission with the aid of a massive MIMO relay was early investigated in [75, 109]. [109] considered secrecy transmission aided by large scale MIMO relaying and provided a thorough secrecy outage capacity-based analysis for two classical types of relay, i.e., AF and DF, respectively, under different channel conditions, based on which asymptotic analysis was further developed to gain insights on the choice between these two relaying schemes. In addition, to the best knowledge of the author, secrecy cooperative transmission jointly considering large scale WEH-enabled MIMO AF relaying was first investigated in [75] and the explicit secrecy outage capacity was obtained therein to analyze the impact of imperfect legitimate CSI and no eavesdropper's CSI.

5.3 System Model

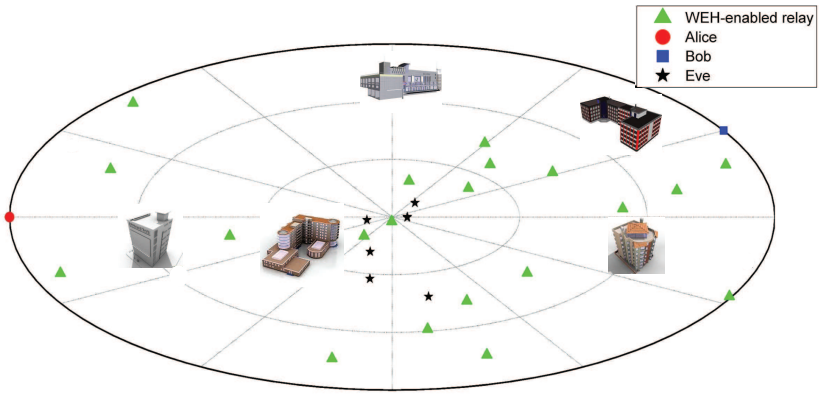


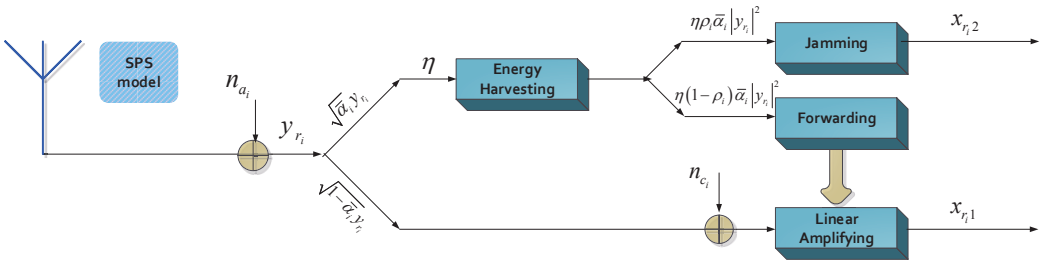
Figure 5.1: The system model for an AF relay-assisted SWIPT WSN.

In this chapter, we consider secrecy transmission in a SWIPT-enabled WSN, where one Tx (Alice) establishes confidential communication with the legitimate Rx (Bob) far away from it assisted by a set of N WEH-enabled sensors (of non-broken communication links from the source and to the destination) working as AF relays,

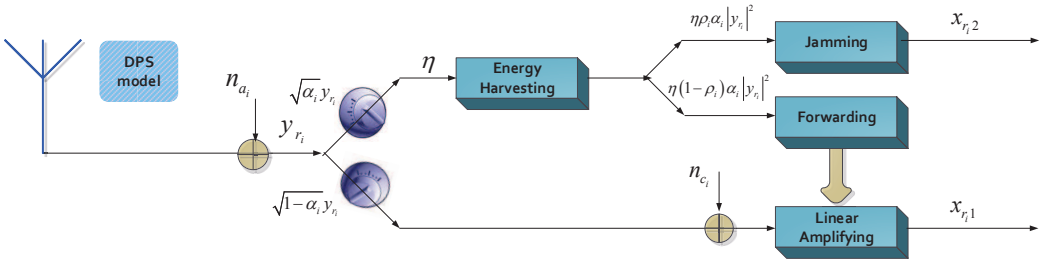
denoted by $\mathcal{N} = \{1, 2, \ldots, N\}$, in the presence of multiple eavesdroppers (Eves), denoted by $\mathcal{K} = \{1, 2, \ldots, K\}$, all equipped with single antenna. Our CB mixed with CJ scheme is applicable to, but not restricted to scenarios, for example, a remote health system where a moving patient needs to report its physical data to a health centre far away from it with the aid of intermediary sensors installed on other patients living in the same community, an environmental monitor system where a thermal sensor is required to send real-time temperature to a data centre and thus solicits help from large number of surrounding sensors that have limited battery levels. It is assumed that there is no direct link from the Tx to the Rx or Eves[1] due to either severe path loss or building-induced shadowing as seen in Fig. 5.1. We assume a two-hop relaying protocol consisting of two equal-time transmit-slot and the duration of one transmit-slot is normalized to be unit one so that the term of energy and power is interchangeable w.r.t. one transmit-slot. In addition, CSI at the associated Tx is assumed to be known perfectly.

For the receiver at each AF relay, we introduce two types of WEH-enabled receiver architecture, namely, *static power splitting (SPS)* (Fig. 5.2(a)) and *dynamic power splitting (DPS)* (Fig. 5.2(b)), both of which allow the relay to harvest energy and receive information from the same received signal. Specifically, as seen from Fig. 5.2, the receiver first splits a portion of α_i, of the received power for EH and the rest $1 - \alpha_i$ for information receiving, $\forall i$. Next, the α_i portion of harvested power is further divided into two streams with the power ratio ρ_i versus $1 - \rho_i$: $\eta\rho_i\alpha_i|y_{r_i}|^2$ used for generating the AN to confound the eavesdroppers and the other $\eta(1-\rho_i)\alpha_i|y_{r_i}|^2$ for amplifying the received signal, where y_{r_i} is the ith element of the received signal $\boldsymbol{y}_r \in \mathbb{C}^{N\times 1}$ and $0 \leq \eta < 1$ denotes the EH efficiency. Note that DPS with adjustable α_i's is currently the most general receiver operation due to the fact that practical circuits cannot directly decode the information from the stream used for EH [9] and SPS is just a special case of DPS with $\alpha_i = \bar{\alpha}_i$, $\forall i$, fixed for the whole

[1]Note that when there exist direct links, by incorporating destination-aided AN in the first transmit-slot (see [73]), our problem formulation and solutions are shown to be applicable without much modification as well.



(a) WEH-enabled relay with static power splitting.



(b) WEH-enabled relay with dynamic power splitting.

Figure 5.2: Architectures of the receiver for WEH-enabled relay.

transmission duration. However, SPS, advocated for its ease of implementation, is introduced separately in the sequel for its simplified relay beamforming design. It is also assumed that any harvested power from the first transmit-slot is not available for cooperation communication in the following transmission phases.

In the first transmit-slot, the received signal at each individual relay can be expressed as

$$y_{r_i} = h_{sr_i}\sqrt{P_s}s + n_{a,i},\ \forall i, \tag{5.1}$$

where the transmit signal s is a circularly symmetric complex Gaussian (CSCG) RV with zero mean and unit variance, denoted by $s \sim \mathcal{CN}(0,1)$. h_{sr_i} denotes the complex channel from the Tx to the ith relay, P_s is the transmit power at the Tx and $n_{a,i}$ is the AWGN introduced by the receiving antenna of the ith relay, denoted by $n_{a,i} \sim \mathcal{CN}(0, \sigma_{n_a}^2)$. As such, the linearly amplified baseband equivalent signal at the output of the ith relay is given by

$$x_{r_i1} = \beta_i(\sqrt{1-\alpha_i}y_{r_i} + n_{c,i}),\ \forall i, \tag{5.2}$$

where β_i is the complex amplifying coefficient, namely, AF relay weight, and $n_{c,i}$ is the noise due to signal conversion from RF band to baseband, denoted by $n_{c,i} \sim \mathcal{CN}(0, \sigma_{n_c}^2)$. Since x_{r_i1} is constrained by the portion of harvested power for forwarding, i.e., $\eta(1-\rho_i)\alpha_i|y_{r_i}|^2$, β_i is accordingly given by

$$\beta_i = \sqrt{\frac{\eta(1-\rho_i)\alpha_i|h_{sr_i}|^2 P_s}{(1-\alpha_i)|h_{sr_i}|^2 P_s + (1-\alpha_i)\sigma_{n_a}^2 + \sigma_{n_c}^2}} e^{j\measuredangle\beta_i}. \tag{5.3}$$

Next, we introduce the CJ scheme coordinately performed by all relays. Denote the CJ signal generated from N relays by $\boldsymbol{x}_{r2} = [x_{r_1 2}, \ldots, x_{r_N 2}]^T$ and define its covariance matrix as $\boldsymbol{S} = \mathbb{E}[\boldsymbol{x}_{r2}\boldsymbol{x}_{r2}^H]$. Then the coordinated CJ transmission can be uniquely determined by the truncated EVD of $\boldsymbol{S}$ given by $\boldsymbol{S} = \tilde{\boldsymbol{V}}\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{V}}^H$, where $\boldsymbol{\Sigma} = diag([\sigma_1, \ldots, \sigma_d])$ is a diagonal matrix with σ_j's denoting all the positive eigenvalues of $\boldsymbol{S}$, and $\boldsymbol{V} \in \mathbb{C}^{N\times d}$ is the precoding matrix satisfying $\boldsymbol{V}^H\boldsymbol{V} = \boldsymbol{I}$. As a result, the CJ signal can be expressed as

$$\boldsymbol{x}_{r2} = \sum_{j=1}^{d} \sqrt{\sigma_j}\boldsymbol{v}_j s_j', \tag{5.4}$$

where $\boldsymbol{v}_j$'s are drawn from the columns of $\boldsymbol{V}$, and s_j''s are *i.i.d.* complex Gaussian variables denoted by $s_j' \sim \mathcal{CN}(0,1)$, which is known as the worst-case distribution for AN [21]. On the other hand, $\mathbb{E}[|x_{r_i 2}|^2] \leq \eta\rho_i\alpha_i|y_{r_i}|^2$, $\forall i$, denotes the power constraint for jamming at the ith relay, which implies that

$$\mathrm{tr}(\boldsymbol{S}\boldsymbol{E}_i) \leq \eta\rho_i\alpha_i P_s|h_{sr_i}|^2, \ \forall i, \tag{5.5}$$

where $\boldsymbol{E}_i$ is a diagonal matrix with its diagonal denoted by $\boldsymbol{e}_i$ (a unit vector with its ith entry equal to 1 and all the other 0).

Note that the CJ scheme we propose is of the most general form. For a special case of $d = 1$, i.e., $\boldsymbol{x}_{r2} = \sqrt{\sigma_1}\boldsymbol{v}_1 s_1'$, each relay transmits a common jamming signal s_1' with their respective weight drawn from $\boldsymbol{v}_1$ [8, 36]. This case is desirable in practice since it reduces the overhead caused by exchanging d CJ beams.

In a summary, the re-transmit signal at the ith relay is given by

$$x_{r_i} = x_{r_i 1} + x_{r_i 2},\ \forall i. \tag{5.6}$$

According to (5.6) together with (5.1), (5.2), and (5.4), the transmit signal from all relays can be expressed as a vector given by

$$\boldsymbol{x}_r = \boldsymbol{D}_{\beta\alpha}\boldsymbol{h}_{sr}\sqrt{P_s}s + \boldsymbol{D}_{\beta\alpha}\boldsymbol{n}_a + \boldsymbol{D}_{\beta}\boldsymbol{n}_c + \sum_{j=1}^{d}\sqrt{\sigma_j}\boldsymbol{v}_j s'_j, \tag{5.7}$$

where$\boldsymbol{D}_{\beta\alpha}$ and $\boldsymbol{D}_{\beta}$ denote diagonal matrices with their diagonals composed of $(\beta_1\sqrt{1-\alpha_1},\ldots,\beta_N\sqrt{1-\alpha_N})^T$ and $(\beta_1,\ldots,\beta_N)^T$, respectively. In addition, $\boldsymbol{h}_{sr} = [h_{sr_i}]_{i=1}^N$, $\boldsymbol{n}_a = [n_{a,i}]_{i=1}^N$, and $\boldsymbol{n}_c = [n_{c,i}]_{i=1}^N$.

In the second transmit-slot, the received signal at the desired receiver, i.e., Bob, is given by

$$y_d = \boldsymbol{h}_{rd}^T\boldsymbol{x}_r + n_d, \tag{5.8}$$

where $\boldsymbol{h}_{rd} = [h_{r_i d}]_{i=1}^N$ comprises complex channels from the ith relay to the Rx, and $n_d \sim \mathcal{CN}(0,\sigma_{n_d}^2)$ is the corresponding receiving AWGN. By substituting (5.7) for $\boldsymbol{x}_r$ in (5.8), y_d can be expressed as

$$y_d = \boldsymbol{h}_{rd}^T\boldsymbol{D}_{\beta\alpha}\boldsymbol{h}_{sr}\sqrt{P_s}s + \boldsymbol{h}_{rd}^T\boldsymbol{D}_{\beta\alpha}\boldsymbol{n}_a + \boldsymbol{h}_{rd}^T\boldsymbol{D}_{\beta}\boldsymbol{n}_c + \boldsymbol{h}_{rd}^T\sum_{j=1}^{d}\sqrt{\sigma_j}\boldsymbol{v}_j s'_j + n_d. \tag{5.9}$$

The received signal at the kth Eve, $\forall k \in \mathcal{K}$, is similarly given by

$$y_{e,k} = \boldsymbol{h}_{re,k}^T\boldsymbol{D}_{\beta\alpha}\boldsymbol{h}_{sr}\sqrt{P_s}s + \boldsymbol{h}_{re,k}^T\boldsymbol{D}_{\beta\alpha}\boldsymbol{n}_a + \boldsymbol{h}_{re,k}^T\boldsymbol{D}_{\beta}\boldsymbol{n}_c + \boldsymbol{h}_{re,k}^T\sum_{j=1}^{d}\sqrt{\sigma_j}\boldsymbol{v}_j s'_j + n_{e,k}, \tag{5.10}$$

where $\boldsymbol{h}_{re,k} = [h_{r_i e,k}]_{i=1}^N$ denotes the complex channels from relays to the kth Eve, and $n_e \sim \mathcal{CN}(0,\sigma_{n_e}^2)$ is the AWGN at the kth eavesdropper.

The instantaneous mutual information for the legitimate Rx is $r_{\mathrm{S,D}} = \frac{1}{2}\log_2(1+$

$\mathrm{SINR}_{\mathrm{S,D}}$), and that for the kth Eve is $r_{\mathrm{S,E,k}} = \frac{1}{2}\log_2(1+\mathrm{SINR}_{\mathrm{S,E,k}})$, $\forall k$, where $\mathrm{SINR}_{\mathrm{S,D}}$ and $\mathrm{SINR}_{\mathrm{S,E,k}}$ denote their respective SINR as follows.

$$\mathrm{SINR}_{\mathrm{S,D}} = \frac{P_s|\boldsymbol{h}_{rd}^T\boldsymbol{D}_{\beta\alpha}\boldsymbol{h}_{sr}|^2}{\mathrm{tr}(\boldsymbol{S}\boldsymbol{h}_{rd}^\dagger\boldsymbol{h}_{rd}^T)+\sigma_{n_a}^2\|\boldsymbol{h}_{rd}^T\boldsymbol{D}_{\beta\alpha}\|^2+\sigma_{n_c}^2\|\boldsymbol{h}_{rd}^T\boldsymbol{D}_{\beta}\|^2+\sigma_{n_d}^2} \tag{5.11}$$

$$\mathrm{SINR}_{\mathrm{S,E,k}} = \frac{P_s|\boldsymbol{h}_{re,k}^T\boldsymbol{D}_{\beta\alpha}\boldsymbol{h}_{sr}|^2}{\mathrm{tr}(\boldsymbol{S}\boldsymbol{h}_{re,k}^\dagger\boldsymbol{h}_{re,k}^T)+\sigma_{n_a}^2\|\boldsymbol{h}_{re,k}^T\boldsymbol{D}_{\beta\alpha}\|^2+\sigma_{n_c}^2\|\boldsymbol{h}_{re,k}^T\boldsymbol{D}_{\beta}\|^2+\sigma_{n_{e,k}}^2} \tag{5.12}$$

Next, we define the secrecy rate region that consists of all the achievable secrecy rate for the relay wiretap channel given transmit power P_s, denoted by $\mathcal{R}(\{\measuredangle\beta_i\},\{\rho_i\},\{\alpha_i\},\boldsymbol{S})$, which is given by [8, 20]

$$\mathcal{R}(\{\measuredangle\beta_i\},\{\rho_i\},\{\alpha_i\},\boldsymbol{S}) \triangleq \bigcup_{\{\{\measuredangle\beta_i\},\{\rho_i\},\{\alpha_i\}\}} \Big\{ r_{\mathrm{sec}} : r_{\mathrm{sec}} \le \big(r_{\mathrm{S,D}} - \max_{k\in\mathcal{K}}\ r_{\mathrm{S,E,k}}\big)^+,\ (5.5),\ \boldsymbol{S}\succeq\boldsymbol{0}\Big\}. \tag{5.13}$$

5.4 Problem Formulation

5.4.1 AN-Aided Secrecy Relay Beamforming for SPS

In this section, we consider the secrecy rate maximization problem by jointly optimizing the AN beams, relay beam and their power allocations for WEH-enabled AF relays operating with SPS, i.e., $\alpha_i = \bar{\alpha}_i$, $\forall i$, is fixed.

By replacing β_i with (5.3), $|\boldsymbol{h}_{rd}^T\boldsymbol{D}_{\beta\alpha}\boldsymbol{h}_{sr}|^2$ in (5.11) can be expressed as follows.

$$|\boldsymbol{h}_{rd}^T\boldsymbol{D}_{\beta\alpha}\boldsymbol{h}_{sr}|^2 = |\sum_{i=1}^{N} w_{1,i}[\tilde{\boldsymbol{h}}_{sd}]_i|^2, \tag{5.14}$$

where $w_{1,i} = \sqrt{1-\rho_i}e^{j\measuredangle\beta_i}$ and $[\tilde{\boldsymbol{h}}_{sd}]_i$ is defined by

$$h_{sr_i}h_{r_id}\sqrt{\frac{\eta\bar{\alpha}_i(1-\bar{\alpha}_i)|h_{sr_i}|^2P_s}{(1-\bar{\alpha}_i)(|h_{sr_i}|^2P_s+\sigma_{n_a}^2)+\sigma_{n_c}^2}} \tag{5.15}$$

for convenience. In addition, $\sigma_{n_a}^2\|\boldsymbol{h}_{rd}^T\boldsymbol{D}_{\beta\alpha}\|^2$ and $\sigma_{n_c}^2\|\boldsymbol{h}_{rd}^T\boldsymbol{D}_{\beta}\|^2$ in (5.11) can also be

combined as follows.

$$\sigma_{n_a}^2\|\boldsymbol{h}_{rd}^T\boldsymbol{D}_{\beta\alpha}\|^2+\sigma_{n_c}^2\|\boldsymbol{h}_{rd}^T\boldsymbol{D}_{\beta}\|^2=\sum_{i=1}^{N}|w_{1,i}|^2[\boldsymbol{D}_{\hat{sd}}]_{i,i}, \tag{5.16}$$

where

$$[\boldsymbol{D}_{\hat{sd}}]_{i,i}=\frac{\eta\bar{\alpha}_iP_s|h_{sr_i}|^2|h_{r_id}|^2((1-\bar{\alpha}_i)\sigma_{n_a}^2+\sigma_{n_c}^2)}{(1-\bar{\alpha}_i)(|h_{sr_i}|^2P_s+\sigma_{n_a}^2)+\sigma_{n_c}^2}. \tag{5.17}$$

As a result, $r_{\mathrm{S,D}}$ can be rewritten as:

$$r_{\mathrm{S,D}}=\frac{1}{2}\log_2\left(1+\frac{P_s|\tilde{\boldsymbol{h}}_{sd}^T\boldsymbol{w}_1|^2}{\mathrm{tr}(\boldsymbol{S}\boldsymbol{h}_{rd}^{\dagger}\boldsymbol{h}_{rd}^T)+\boldsymbol{w}_1^H\boldsymbol{D}_{\hat{sd}}\boldsymbol{w}_1+\sigma_{n_d}^2}\right), \tag{5.18}$$

where $\boldsymbol{w}_1=[w_{1,i}]_{i=1}^N$. Similarly by denoting

$$h_{sr_i}h_{r_ie,k}\sqrt{\frac{\eta\bar{\alpha}_i(1-\bar{\alpha}_i)|h_{sr_i}|^2P_s}{(1-\bar{\alpha}_i)(|h_{sr_i}|^2P_s+\sigma_{n_a}^2)+\sigma_{n_c}^2}} \tag{5.19}$$

as $[\tilde{\boldsymbol{h}}_{se,k}]_i$ and

$$\frac{\eta\bar{\alpha}_iP_s|h_{sr_i}|^2|h_{r_ie,k}|^2((1-\bar{\alpha}_i)\sigma_{n_a}^2+\sigma_{n_c}^2)}{(1-\bar{\alpha}_i)(|h_{sr_i}|^2P_s+\sigma_{n_a}^2)+\sigma_{n_c}^2} \tag{5.20}$$

as $[\boldsymbol{D}_{\hat{se},k}]_{i,i}$, $\forall i\in\mathcal{N}$, $\forall k\in\mathcal{K}$, $r_{\mathrm{S,E,k}}$ is simplified as:

$$r_{\mathrm{S,E,k}}=\frac{1}{2}\log_2\left(1+\frac{P_s|\tilde{\boldsymbol{h}}_{se,k}^T\boldsymbol{w}_1|^2}{\mathrm{tr}(\boldsymbol{S}\boldsymbol{h}_{re,k}^{\dagger}\boldsymbol{h}_{re,k}^T)+\boldsymbol{w}_1^H\boldsymbol{D}_{\hat{se},k}\boldsymbol{w}_1+\sigma_{n_e,k}^2}\right). \tag{5.21}$$

By some simple manipulation, (5.5) is reformulated as a power per-relay jamming power constraint given by

$$\mathrm{tr}(\boldsymbol{S}\boldsymbol{E}_i)\leq\eta\bar{\alpha}_iP_s|h_{sr_i}|^2(1-|w_{1,i}|^2),\ \forall i. \tag{5.22}$$

Now, the secrecy rate maximization problem w.r.t. ρ_i's, $\measuredangle\beta_i$'s and $\boldsymbol{S}$ for

SPS-based relays can be formulated as below.

$$(\text{P1}): \max_{\boldsymbol{w}_1, \boldsymbol{S}} \ ((5.18) - \max_{k\in\mathcal{K}} (5.21))^+ \ \ \texttt{s.t.}\ (5.22),\ \boldsymbol{S} \succeq \boldsymbol{0}.$$

5.4.2 AN-Aided Secrecy Relay Beamforming for DPS

In this section, we consider the secrecy rate maximization problem for WEH-enabled AF relays with adjustable power splitting ratios $\{\alpha_i\}$ by jointly optimizing the AN beams and relay beam, WEH power splitting ratios $\{\alpha_i\}$, AN power splitting ratios $\{\rho_i\}$ and amplifying phases $\{\measuredangle\beta_i\}$.

First, consider the following variable transformation:

$$\begin{cases} u_{1,i} = \sqrt{\frac{\alpha_i(1-\alpha_i)(1-\rho_i)}{(1-\alpha_i)(|h_{sr_i}|^2P_s+\sigma_{n_a}^2)+\sigma_{n_c}^2}} e^{j\measuredangle\beta_i} \\ u_{2,i} = \sqrt{\frac{\alpha_i(1-\rho_i)}{(1-\alpha_i)(|h_{sr_i}|^2P_s+\sigma_{n_a}^2)+\sigma_{n_c}^2}} \end{cases}, \ \forall i, \tag{5.23}$$

via which, $|\boldsymbol{h}_{rd}^T\boldsymbol{D}_{\beta\alpha}\boldsymbol{h}_{sr}|^2$ can be expressed as $|\boldsymbol{s}_{sd}^T\boldsymbol{u}_1|^2$, where $\boldsymbol{s}_{sd} = [h_{sr_i}h_{r_id}\sqrt{\eta|h_{sr_i}|^2P_s}]_{i=1}^N$, $\boldsymbol{u}_1 = [u_{1,i}]_{i=1}^N$. Moreover, $\|\boldsymbol{h}_{rd}^T\boldsymbol{D}_{\beta\alpha}\|^2$ and $\|\boldsymbol{h}_{rd}^T\boldsymbol{D}_{\beta}\|^2$ can be simplified as $\boldsymbol{u}_1^H diag(\boldsymbol{c}_0 \circ \|\boldsymbol{h}_{rd}\|.^2)\boldsymbol{u}_1$ and $\boldsymbol{u}_2^H diag(\boldsymbol{c}_0 \circ \|\boldsymbol{h}_{rd}\|.^2)\boldsymbol{u}_2$, respectively, where $\boldsymbol{c}_0 = [c_{0,i}]_{i=1}^N$ with $c_{0,i} = \eta P_s|h_{sr_i}|^2$, $\forall i$, and $\boldsymbol{u}_2 = [u_{2,i}]_{i=1}^N$. Similarly, we have $|\boldsymbol{h}_{re,k}^T\boldsymbol{D}_{\beta\alpha}\boldsymbol{h}_{sr}|^2 = |\boldsymbol{s}_{se,k}^T\boldsymbol{u}_1|^2$, $\|\boldsymbol{h}_{re,k}^T\boldsymbol{D}_{\beta\alpha}\|^2 = \boldsymbol{u}_1^H diag(\boldsymbol{c}_0 \circ \|\boldsymbol{h}_{re,k}\|.^2)\boldsymbol{u}_1$ and $\|\boldsymbol{h}_{re,k}^T\boldsymbol{D}_{\beta}\|^2 = \boldsymbol{u}_2^H diag(\boldsymbol{c}_0 \circ \|\boldsymbol{h}_{re,k}\|.^2)\boldsymbol{u}_2$, where $\boldsymbol{s}_{se,k} = [h_{sr_i}h_{r_ie,k}\sqrt{\eta|h_{sr_i}|^2P_s}]_{i=1}^N$, $\forall k \in \mathcal{K}$.

Second, apply the above transformation to $\text{SINR}_{\text{S,D}}$ (c.f. (5.11)) and $\text{SINR}_{\text{S,E,k}}$ (c.f. (5.12)), $\forall k$, and return (5.24) and (5.25) as follows.

$$\text{SINR}_{\text{S,D}} = \frac{P_s|\boldsymbol{s}_{sd}^T\boldsymbol{u}_1|^2}{\text{tr}(\boldsymbol{S}\boldsymbol{h}_{rd}^\dagger\boldsymbol{h}_{rd}^T)+\sigma_{n_a}^2\boldsymbol{u}_1^H diag(\boldsymbol{c}_0\circ\|\boldsymbol{h}_{rd}\|.^2)\boldsymbol{u}_1+\sigma_{n_c}^2\boldsymbol{u}_2^H diag(\boldsymbol{c}_0\circ\|\boldsymbol{h}_{rd}\|.^2)\boldsymbol{u}_2+\sigma_{n_d}^2} \tag{5.24}$$

$$\text{SINR}_{\text{S,E,k}} = \frac{P_s|\boldsymbol{s}_{se,k}^T\boldsymbol{u}_1|^2}{\text{tr}(\boldsymbol{S}\boldsymbol{h}_{re,k}^\dagger\boldsymbol{h}_{re,k}^T)+\sigma_{n_a}^2\boldsymbol{u}_1^H diag(\boldsymbol{c}_0\circ\|\boldsymbol{h}_{re,k}\|.^2)\boldsymbol{u}_1+\sigma_{n_c}^2\boldsymbol{u}_2^H diag(\boldsymbol{c}_0\circ\|\boldsymbol{h}_{re,k}\|.^2)\boldsymbol{u}_2+\sigma_{n_{e,k}}^2} \tag{5.25}$$

Next, we recast constraints w.r.t. $\boldsymbol{S}$, α_i's and ρ_i's to those w.r.t. the transformed variables $u_{1,i}$'s and $u_{2,i}$'s. In accordance with (5.23), the optimization variables, α_i's

and ρ_i's, can be alternatively given by

$$\begin{cases} \alpha_i = 1 - \frac{|u_{1,i}|^2}{|u_{2,i}|^2} \\ \rho_i = 1 - \frac{|u_{2,i}|^2(c_{1,i}|u_{1,i}|^2+\sigma_{n_c}^2|u_{2,i}|^2)}{|u_{2,i}|^2-|u_{1,i}|^2} \end{cases}, \ \forall i, \tag{5.26}$$

where $c_{1,i} = P_s|h_{sr_i}|^2 + \sigma_{n_a}^2$. Replacing α_i's and ρ_i's with (5.26), (5.5) is reformulated as

$$\mathrm{tr}(\boldsymbol{S}\boldsymbol{E}_i) \leq c_{0,i}\left(1 - \frac{|u_{2,i}|^2(c_{1,i}|u_{1,i}|^2+\sigma_{n_c}^2|u_{2,i}|^2)}{|u_{2,i}|^2-|u_{1,i}|^2}\right)\left(1 - \frac{|u_{1,i}|^2}{|u_{2,i}|^2}\right), \ \forall i. \tag{5.27}$$

On the other hand, since $\alpha_i \geq 0$ and $\rho_i \geq 0$, $\forall i$, after some simple manipulation, it follows from (5.26) that

$$|u_{1,i}|^2 - |u_{2,i}|^2 \leq 0, \ \forall i, \tag{5.28}$$

$$|u_{2,i}|^2(c_{1,i}|u_{1,i}|^2 + \sigma_{n_c}^2|u_{2,i}|^2) \leq |u_{2,i}|^2 - |u_{1,i}|^2, \ \forall i. \tag{5.29}$$

Consequently, we are ready to state the secrecy rate maximization problem jointly w.r.t. α_i's, ρ_i's, $\measuredangle\beta_i$'s and $\boldsymbol{S}$ for DPS-based relays as follows.

$$\begin{aligned} (\mathrm{P2}): \max_{\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{S}} \quad & \left(\frac{1}{2}\log_2(1+\mathrm{SINR}_{\mathrm{S,D}}) - \frac{1}{2}\log_2(1+\max_{k\in\mathcal{K}}\mathrm{SINR}_{\mathrm{S,E,k}})\right)^+ \\ \mathtt{s.t.} \quad & (5.27),\ (5.28),\ \text{and}\ (5.29). \end{aligned}$$

5.5 Secure Multi-AF Relaying: A Centralized Approach

In this section, we resort to centralized approaches to solve problem (P1) and (P2), respectively, assuming that there is a coordinating centre that is able to collect all CSIs including $\boldsymbol{h}_{sr}$, $\boldsymbol{h}_{rd}$ and $\boldsymbol{h}_{re}$ perfectly. On the ground of these information, the centre performs optimization and broadcasts to relays their individual optimized parameters, such as the power splitting ratios (α_i's and/ or ρ_i's), the phase of the

complex amplifying coefficient ($\measuredangle\beta_i$'s), and their respective jamming signal $x_{r_i 2}$'s.

5.5.1 Optimal Solutions for SPS

Note that one of the main challenges for solving (P1) lies in the fact that $r_{\mathrm{S,D}}$ and $r_{\mathrm{S,E,k}}$ share the similar structure, which sabotages the convexity of the objective function even if $r_{\mathrm{S,D}}$ and $r_{\mathrm{S,E}}$ are made convex, respectively. Hence, in the following, we recast (P1) into a two-stage problem by introducing a slack variable τ as follows. First, solve the epigraph reformulation of (P1) with a fixed $\tau \in (0, 1]$ as follows.

$$\begin{aligned}(\text{P1.1}): \max_{\boldsymbol{w}_1, \boldsymbol{S}} \ & \frac{P_s|\tilde{\boldsymbol{h}}_{sd}^T \boldsymbol{w}_1|^2}{\mathrm{tr}(\boldsymbol{S}\boldsymbol{h}_{rd}^\dagger \boldsymbol{h}_{rd}^T)+\boldsymbol{w}_1^H \boldsymbol{D}_{\hat{sd}} \boldsymbol{w}_1+\sigma_{n_d}^2} \\ \mathtt{s.t.}\ & 1+\frac{P_s|\tilde{\boldsymbol{h}}_{se,k}^T \boldsymbol{w}_1|^2}{\mathrm{tr}(\boldsymbol{S}\boldsymbol{h}_{re,k}^\dagger \boldsymbol{h}_{re,k}^T)+\boldsymbol{w}_1^H \boldsymbol{D}_{\hat{se},k} \boldsymbol{w}_1+\sigma_{n_e,k}^2} \le 1/\tau, \ \forall k, \\ & (5.22), \ \boldsymbol{S} \succeq \boldsymbol{0}.\end{aligned}$$

Next, define $f_1(\tau)$ as the optimum value of problem (P1.1) and denote $H_1(\tau) = \tau f_1(\tau)$, the objective function of (P1) can be rewritten as

$$\frac{1}{2}\log_2(1+f_1(\tau)) - \frac{1}{2}\log_2(1/\tau) = \frac{1}{2}\log_2(\tau + H_1(\tau)), \tag{5.30}$$

in which $(\cdot)^+$ in the objective function of (P1) has been omitted and we claim a zero secrecy rate if (5.30) admits a negative value. As a result, problem (P1) can be equivalently given by

$$\begin{aligned}(\text{P1.2}): \max_{\tau} \ & \frac{1}{2}\log_2(\tau + H_1(\tau)) \\ \mathtt{s.t.}\ & \tau_{\min,1} \le \tau \le 1,\end{aligned}$$

which is noteworthy since this single-variable optimization problem allows for simple one-dimension search algorithm over $\tau \in [\tau_{\min,1}, 1]$, assuming that $H_1(\tau)$ is attainable given any τ in this region. As the physical meaning of $1/\tau - 1$ in (P1.1) can be interpreted as the maximum permitted SINR for the best eavesdropper's channel,

feasibility for a non-zero secrecy rate implies that

$$\begin{aligned}
&\frac{P_s|\tilde{\boldsymbol{h}}_{sd}^T\boldsymbol{w}_1|^2}{\mathrm{tr}(\boldsymbol{S}\boldsymbol{h}_{rd}^\dagger\boldsymbol{h}_{rd}^T)+\boldsymbol{w}_1^H\boldsymbol{D}_{\hat{sd}}\boldsymbol{w}_1+\sigma_{n_d}^2} \geq 1/\tau - 1 \\
\Leftrightarrow\ \tau &\geq \frac{1}{1+\frac{P_s|\tilde{\boldsymbol{h}}_{sd}{}^T\boldsymbol{w}_1|^2/\sigma_{n_d}^2}{\mathrm{tr}(\boldsymbol{S}\boldsymbol{h}_{rd}^\dagger\boldsymbol{h}_{rd}^T)/\sigma_{n_d}^2+\boldsymbol{w}_1^H\boldsymbol{D}_{\hat{sd}}\boldsymbol{w}_1/\sigma_{n_d}^2+1}} \\
&\overset{(a)}{\geq} \frac{1}{1+P_s\|\tilde{\boldsymbol{h}}_{sd}\|^2\|\boldsymbol{w}_1\|^2/\sigma_{n_d}^2} \\
&\overset{(b)}{\geq} \frac{1}{1+NP_s\|\tilde{\boldsymbol{h}}_{sd}\|^2/\sigma_{n_d}^2} = \tau_{\min,1},
\end{aligned} \tag{5.31}$$

where Cauchy-Schwarz inequality has been applied in (a) and (b) follows from $|w_{1,i}|^2 \leq 1$, $\forall i \in \mathcal{N}$.

Note that the above epigraph reformulation of non-convex problems like (P1) has been widely employed in literature [33, 51], and therefore a detailed proof is thus skipped here that problem (P1.2) admits the same optimum value as problem (P1) while problem (P1.1) with the optimal τ provides the corresponding optimal solution to (P1). As a result, we summarize the steps solving (P1): given any $\tau \in [\tau_{\min,1}, 1]$, solve (P1.1) to obtain $H_1(\tau)$; solve (P1.2) via a one-dimension search over τ. Before developing solutions to (P1.1), we highlight the lemma as below.

Lemma 5.5.1. *$H_1(\tau)$ is a concave function of τ.*

Proof. See Appendix G. □

Remark 5.5.1. *Thanks to Lemma 5.5.1, it is easy to verify that $\frac{1}{2}\log_2(\tau + H_1(\tau))$ is also a concave function of τ according to the composition rule [84, pp. 84], which allows for a more effective one-dimension search for the optimum τ, for example, bi-section method, than exhaustive search used in [57]. On the other hand, other one-dimension search methods such as coordinate search and golden search etc. have been employed to solve this problem in [51] though, they only yield local optimum solutions in general without showing Lemma 5.5.1. Moreover, although $H_1(\tau)$ is not derivable w.r.t. τ, the bi-section method can be numerically implemented derivative-free, which is shown in detail in Table 5.1.*

In the sequel we focus on solving (P1.1). By introducing $\boldsymbol{X}_1 = \boldsymbol{w}_1\boldsymbol{w}_1^H$ and ignoring the rank-one constraint on $\boldsymbol{X}_1$, (P1.1) can be alternatively solved by the following problem.

$$\text{(P1.1-SDR)}: \begin{cases} \underset{\boldsymbol{X}_1,\boldsymbol{S}}{\max} \ \frac{\tau P_s \mathrm{tr}(\boldsymbol{X}_1 \tilde{\boldsymbol{h}}_{sd}^{\dagger} \tilde{\boldsymbol{h}}_{sd}^{T})}{\mathrm{tr}(\boldsymbol{S}\boldsymbol{h}_{rd}^{\dagger}\boldsymbol{h}_{rd}^{T})+\mathrm{tr}(\boldsymbol{X}_1\boldsymbol{D}_{\hat{sd}})+\sigma_{n_d}^2} \\ \mathtt{s.t.} \ \frac{P_s \mathrm{tr}(\boldsymbol{X}_1 \tilde{\boldsymbol{h}}_{se,k}^{\dagger} \tilde{\boldsymbol{h}}_{se,k}^{T})}{\mathrm{tr}(\boldsymbol{S}\boldsymbol{h}_{re,k}^{\dagger}\boldsymbol{h}_{re,k}^{T})+\mathrm{tr}(\boldsymbol{X}_1\boldsymbol{D}_{\hat{se},k})+\sigma_{n_e,k}^2} \le \frac{1}{\tau} - 1, \ \forall k, \\ \quad \mathrm{tr}((\boldsymbol{S} + \eta\bar{\alpha}_i P_s |h_{sr_i}|^2 \boldsymbol{X}_1)\boldsymbol{E}_i) \le \eta\bar{\alpha}_i P_s |h_{sr_i}|^2, \ \forall i, \\ \quad \boldsymbol{X}_1 \succeq \boldsymbol{0}, \ \boldsymbol{S} \succeq \boldsymbol{0}. \end{cases}$$

Note that the objective function has been multiplied by τ compared with that of (P1.1) in order for the straightforward computation of $H_1(\tau)$.

Although (P1.1-SDR) is made easier to solve than the original (P1.1) by rank relaxation, it is still a quasi-convex problem considering the linear fractional form objective function and constraint, for which, Charnes-Cooper transformation [92] turns out to be an effective tool in equivalent convex reformulation. Specifically, by substituting $\boldsymbol{X}_1 = \hat{\boldsymbol{X}}_1/\xi$ and $\boldsymbol{S} = \hat{\boldsymbol{S}}/\xi$ into problem (P1.1-SDR), it follows that

$$\text{(P1.1-SDP)}: \begin{cases} \underset{\hat{\boldsymbol{X}}_1,\hat{\boldsymbol{S}},\xi\ge 0}{\max} \quad P_s \mathrm{tr}(\hat{\boldsymbol{X}}_1 \tilde{\boldsymbol{h}}_{sd}^{\dagger} \tilde{\boldsymbol{h}}_{sd}^{T}) \\ \quad \mathtt{s.t.} \ \mathrm{tr}(\hat{\boldsymbol{S}}\boldsymbol{h}_{rd}^{\dagger}\boldsymbol{h}_{rd}^{T}) + \mathrm{tr}(\hat{\boldsymbol{X}}_1\boldsymbol{D}_{\hat{sd}}) + \xi\sigma_{n_d}^2 = \tau, \\ \quad\quad (\frac{1}{\tau} - 1)\left(\mathrm{tr}(\hat{\boldsymbol{S}}\boldsymbol{h}_{re,k}^{\dagger}\boldsymbol{h}_{re,k}^{T}) + \mathrm{tr}(\hat{\boldsymbol{X}}_1\boldsymbol{D}_{\hat{se},k}) + \xi\sigma_{n_e,k}^2\right) \ge P_s \mathrm{tr}(\hat{\boldsymbol{X}}_1 \tilde{\boldsymbol{h}}_{se,k}^{\dagger} \tilde{\boldsymbol{h}}_{se,k}^{T}), \ \forall k, \\ \quad\quad \mathrm{tr}((\hat{\boldsymbol{S}} + \eta\bar{\alpha}_i P_s |h_{sr_i}|^2 \hat{\boldsymbol{X}}_1)\boldsymbol{E}_i) \le \xi\eta\bar{\alpha}_i P_s |h_{sr_i}|^2, \ \forall i, \\ \quad\quad \hat{\boldsymbol{X}}_1 \succeq \boldsymbol{0}, \ \hat{\boldsymbol{S}} \succeq \boldsymbol{0}. \end{cases}$$

Problem (P1.1-SDP) can now be efficiently solved using interior-point based methods by some off-the-shelf convex optimization toolboxes, e.g., CVX [110].

A natural question comes out that whether the solution derived to (P1.1-SDR)

is also optimal to (P1.1), since the optimum value of problem (P1.1-SDR) only serves as an upper-bound for problem (P1.1) in general due to the rank relaxation. The following proposition answers this question.

Proposition 5.5.1. *1) The optimal solution to problem* (P1.1-SDP) *satisfies* $\mathrm{rank}(\hat{\boldsymbol{X}}_1^*) = 1$*;*

2) $\hat{\boldsymbol{X}}_1^* = \hat{\boldsymbol{w}}_1^* \hat{\boldsymbol{w}}_1^{*H}$*, in which* $\hat{\boldsymbol{w}}_1^*$ *is given by*

$$\hat{\boldsymbol{w}}_1^* = \sqrt{\frac{\tau - \xi^* \sigma_{n_d}^2 - \mathrm{tr}(\hat{\boldsymbol{S}}^* \boldsymbol{h}_{rd}^\dagger \boldsymbol{h}_{rd}^T)}{\mathrm{tr}(\hat{\boldsymbol{w}}_1 \hat{\boldsymbol{w}}_1^H \boldsymbol{D}_{\hat{s}d})}} \hat{\boldsymbol{w}}_1, \tag{5.32}$$

where $\hat{\boldsymbol{w}}_1$ *is given in Appendix H (c.f.* (H.9)*);*

3) $\mathrm{rank}(\hat{\boldsymbol{S}}^*) \leq \min(K, N)$.

Proof. See Appendix H. □

Corollary 5.5.1. *When* $N = 1$*,* $K = 1$*,* $\hat{s}^* = 0$.

Proof. When $N = 1$, $K = 1$, the objective function of (P1) reduces to (c.f. (5.18) and (5.21))

$$\left(\frac{1}{2} \log_2 \left(1 + \frac{P_s |\tilde{h}_{sd}|^2 |w_1|^2}{|h_{rd}|^2 s + D_{\hat{s}d} |w_1|^2 + \sigma_{n_d}^2} \right) - \frac{1}{2} \log_2 \left(1 + \frac{P_s |\tilde{h}_{se}|^2 |w_1|^2}{|h_{re}|^2 s + D_{\hat{s}e} |w_1|^2 + \sigma_{n_e}^2} \right) \right)^+, \tag{5.33}$$

Assuming that problem (P1) is feasible, the first derivative of (5.33) w.r.t. s is given by

$$\frac{\partial (5.33)}{\partial s} = -\frac{n_B |h_{rd}|^2}{(1+R_B) d_B^2} + \frac{n_E |h_{re}|^2}{(1+R_E) d_E^2}, \tag{5.34}$$

where $R_B = \frac{P_s |\tilde{h}_{sd}|^2 |w_1|^2}{|h_{rd}|^2 s + D_{\hat{s}d} |w_1|^2 + \sigma_{n_d}^2}$ with its numerator and denominator denoted by n_B and d_B, respectively, and $R_E = \frac{P_s |\tilde{h}_{se}|^2 |w_1|^2}{|h_{re}|^2 s + D_{\hat{s}e} |w_1|^2 + \sigma_{n_e}^2}$ with n_E and d_E similarly defined. As a result of feasibility, a non-negative secrecy rate indicates that $\frac{n_B}{d_B} \geq \frac{n_E}{d_E}$, which further implies that $\frac{d_B}{n_B} \leq \frac{d_E}{n_E}$ and $\frac{d_B + n_B}{n_B} \leq \frac{d_E + n_E}{n_E}$, and therefore

$\left(\frac{n_B|h_{rd}|^2}{(1+R_B)d_B^2}\right) / \left(\frac{n_E|h_{re}|^2}{(1+R_E)d_E^2}\right) = (\frac{d_E}{n_E}\frac{d_E+n_E}{n_E})/(\frac{d_B}{n_B}\frac{d_B+n_B}{n_B}) \geq 1$. According to (5.34), a non-positive derivative w.r.t. s is returned, which shows that the secrecy rate in this case monotonically decreases with the jamming variance and thus completes the proof. □

Proposition 5.5.1 implies that the rank-one relaxation of (P1.1-SDR) from (P1.1) is tight for an arbitrary given τ. The ρ^*'s and $\measuredangle\beta_i^*$'s can thus be retrieved from the magnitude and angle of $\boldsymbol{w}_1^*$, respectively, by applying EVD to $\boldsymbol{X}_1^*$.

5.5.2 Proposed Solutions for DPS

Similar to Section 5.5.1, in this section, we aim at solving the two-stage reformulation of problem (P2) by introducing a slack variable $\tau \in [\tau_{\min,2}, 1]$. First, for a given τ, we solve the following problem.

$$\begin{aligned}(\text{P2.1}): \max_{\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{S}} \quad & (5.24)\\ \texttt{s.t.} \quad & 1 + (5.25) \leq 1/\tau,\ \forall k,\ (5.27),\ (5.28),\ (5.29).\end{aligned}$$

Next, alike problem (P1.2), denoting $H_2(\tau) = \tau f_2(\tau)$, where $f_2(\tau)$ is the optimum value for problem (P2.1), we solve the following problem that admits the same optimum value as (P2).

$$\begin{aligned}(\text{P2.2}): \max_{\tau} \quad & \frac{1}{2}\log_2(\tau + H_2(\tau)),\\ \texttt{s.t.} \quad & \tau_{\min,2} \leq \tau \leq 1.\end{aligned}$$

$\tau_{\min,2}$ is similarly derived as $\tau_{\min,1}$, so that we directly arrive at $\tau \geq \frac{1}{1+P_s\|\boldsymbol{s}_{sd}\|^2 \sum_{i=1}^{N} \frac{1}{\sigma_{n_d}^2(|h_{sr_i}|^2 P_s+\sigma_{n_a}^2+\sigma_{n_c}^2)}}$, denoted by $\tau_{\min,2}$. We claim that (P2.2) can be solved by bi-section for τ over the interval $[\tau_{\min,2}, 1]$ assuming that $H_2(\tau)$ is valid for any given τ (Otherwise a zero secrecy rate, i.e., $H_2(\tau) = 0$, is returned.), since $H_2(\tau)$ has the following property.

Lemma 5.5.2. *$H_2(\tau)$ is a concave function of τ.*

Proof. The proof is similar to that for Lemma 5.5.1 and thus omitted here for brevity. □

It is also seen that how to attain $H_2(\tau)$ forms the main thrust for solving (P2). However, the constraints in (5.27), (5.28) and (5.29) are not convex w.r.t. $u_{1,i}$ and/or $u_{2,i}$, $\forall i$, due to their high orders and multiplicative structures. (P2.1) thus turns out to be very hard to solve in general. To cope with these non-convex constraints, we introduce the following lemma.

Lemma 5.5.3 ([100])**.** *The restricted hyperbolic constraints which have the form $\boldsymbol{x}^H\boldsymbol{x} \leq yz$, where $\boldsymbol{x} \in \mathbb{C}^{N\times 1}$, $y,\ z \geq 0$, are equivalent to rotated second-order cone (SOC) constraints given by*

$$\left\| \begin{pmatrix} 2\boldsymbol{x} \\ y - z \end{pmatrix} \right\| \leq y + z. \tag{5.35}$$

Proof.

$$\begin{aligned} & \boldsymbol{x}^H\boldsymbol{x} \leq yz \\ \Leftrightarrow\ & y^2 + z^2 + 4\boldsymbol{x}^H\boldsymbol{x} \leq 4yz + y^2 + z^2 \\ \Leftrightarrow\ & (y-z)^2 + \|2\boldsymbol{x}\|^2 \leq (y+z)^2 \overset{(a)}{\Leftrightarrow} (5.35), \end{aligned}$$

where (a) holds true since $y,\ z \geq 0$. □

For convenience, denoting $|u_{1,i}|^2$, $|u_{2,i}|^2$, $\mathrm{tr}(\boldsymbol{S}\boldsymbol{E}_i)$ by x_i, y_i and z_i, respectively,

$\forall i$, (5.27) can be rewritten as follows.

$$\begin{aligned} & z_i \leq c_{0,i}\left(1-\frac{y_i(c_{1,i}x_i+\sigma_{n_c}^2 y_i)}{y_i-x_i}\right)\left(1-\frac{x_i}{y_i}\right) \\ \Leftrightarrow & \frac{z_i}{c_{0,i}} \leq 1-\frac{x_i}{y_i}-(c_{1,i}x_i+\sigma_{n_c}^2 y_i) \\ \Leftrightarrow & (\sigma_{n_c}y_i)^2+\left(\sqrt{\left(1-\tfrac{z_i}{c_{0,i}}\right)\tfrac{1}{c_{1,i}}}\right)^2 \leq \left(1-\tfrac{z_i}{c_{0,i}}-c_{1,sr,i}x_i\right)\left(y_i+\tfrac{1}{c_{1,i}}\right) \end{aligned} \tag{5.36}$$

According to (5.5) and (5.23), it is easily verified that $1-\frac{z_i}{c_{0,sr,i}}-c_{1,sr,i}x_i > 1-\rho_i\alpha_i-(1-\rho_i)\alpha_i \geq 0$. Hence, (5.36) is eligible for Lemma 5.5.3, which is reformulated into the following SOC constraint.

$$\left\| \begin{array}{c} 2\sigma_{n_c}y_i \\ 2\sqrt{\left(1-\frac{z_i}{c_{0,i}}\right)\frac{1}{c_{1,i}}} \\ \left(1-\frac{z_i}{c_{0,i}}-c_{1,i}x_i\right)-\left(y_i+\frac{1}{c_{1,i}}\right) \end{array} \right\| \leq \left(1-\tfrac{z_i}{c_{0,i}}-c_{1,i}x_i\right)+\left(y_i+\tfrac{1}{c_{1,i}}\right) \tag{5.37}$$

Similarly, (5.29) can be simplified as $y_i(c_{1,i}x_i+\sigma_{n_c}^2 y_i) \leq y_i - x_i$, and after some manipulation, it is recast into a constraint of the restricted hyperbolic form as follows.

$$(\sigma_{n_c}y_i)^2+\left(\sqrt{\tfrac{1}{c_{1,i}}}\right)^2 \leq (1-c_{1,i}x_i)\left(y_i+\tfrac{1}{c_{1,i}}\right), \tag{5.38}$$

which is thus, in line with Lemma 5.5.3, equivalent to a SOC constraint given by

$$\left\| \begin{array}{c} 2\sigma_{n_c}y_i \\ 2\sqrt{\frac{1}{c_{1,i}}} \\ (1-c_{1,i}x_i)-\left(y_i+\frac{1}{c_{1,i}}\right) \end{array} \right\| \leq (1-c_{1,i}x_i)+\left(y_i+\tfrac{1}{c_{1,i}}\right). \tag{5.39}$$

At last, (5.28) is apparently a linear constraint w.r.t. x_i and y_i given by

$$x_i - y_i \leq 0, \ \forall i. \tag{5.40}$$

Note that (5.27), (5.29), and (5.28) have so far been equivalently transformed

into the SOC constraints (5.37), (5.39), and the linear constraint (5.40), the latter two of which are jointly convex w.r.t. x_i and y_i, $\forall i$. However, (5.37) is not convex w.r.t. z_i, $\forall i$, yet. To circumvent this, in the sequel we propose to solve problem (P2) by alternating optimization. The upshot of the algorithm is that first we fix $\boldsymbol{S}$ by $\overline{\boldsymbol{S}}$ and z_i by $\bar{z}_i = \mathrm{tr}(\overline{\boldsymbol{S}}\boldsymbol{E}_i)$, $\forall i$, and solve problem (P2′)[2] to find optimal $\{\alpha^*\}$, $\{\rho^*\}$ and $\{\measuredangle\beta_i\}$ via (P2′.1) and (P2′.2); then with $\bar{\alpha}_i = \alpha_i^*$, $\forall i$, provided, we devise the optimal solution derived in Section 5.5.1 to obtain the optimal CJ covariance, viz, $\boldsymbol{S}^*$, and thus $z_i^* = \mathrm{tr}(\boldsymbol{S}^*\boldsymbol{E}_i)$, $\forall i$; finally, by updating $\overline{\boldsymbol{S}} = \boldsymbol{S}^*$ and $\bar{z}_i = z_i^*$, $\forall i$, (P2′) and (P1) are iteratively solved until they converge.

The remaining challenges lie in how to solve problem (P2′.1) now that (5.37), (5.39) and (5.40) are all made convex w.r.t. their variables x_i, y_i, $\forall i$. Similar as that for (P1.1), we introduce $\boldsymbol{U}_1 = \boldsymbol{u}_1\boldsymbol{u}_1^H$ and $\boldsymbol{U}_2 = \boldsymbol{u}_2\boldsymbol{u}_2^H$ and exempt problem (P2′.1) from $\mathrm{rank}(\boldsymbol{U}_1) = 1$ and $\mathrm{rank}(\boldsymbol{U}_2) = 1$ as follows.

$$
\begin{aligned}
&(\text{P2}'.1\text{-SDR}):\\
&\begin{cases}
\underset{\boldsymbol{U}_1,\boldsymbol{U}_2,\{x_i\},\{y_i\}}{\mathtt{max}} & (5.41)\\
\qquad\mathtt{s.t.} & (5.42),\ \forall k,\ (5.37),\ (5.39),\ (5.40),\\
& \mathrm{tr}(\boldsymbol{U}_1\boldsymbol{E}_i) = x_i,\ \mathrm{tr}(\boldsymbol{U}_2\boldsymbol{E}_i) = y_i,\ \forall i,\\
& \boldsymbol{U}_1 \succeq \boldsymbol{0},\ \boldsymbol{U}_2 \succeq \boldsymbol{0},
\end{cases}
\end{aligned}
$$

the objective function of which is shown to be $H_2(\tau)$, and (5.41) and (5.42) are shown as below.

$$
\frac{\tau P_s \mathrm{tr}(\boldsymbol{U}_1\boldsymbol{s}_{sd}^\dagger\boldsymbol{s}_{sd}^T)}{\mathrm{tr}(\overline{\boldsymbol{S}}\boldsymbol{h}_{rd}^\dagger\boldsymbol{h}_{rd}^T)+\mathrm{tr}((\sigma_{n_a}^2\boldsymbol{U}_1+\sigma_{n_c}^2\boldsymbol{U}_2)diag(\boldsymbol{c}_0\circ\|\boldsymbol{h}_{rd}\|^{.2}))+\sigma_{n_d}^2} \tag{5.41}
$$

$$
1+\frac{P_s \mathrm{tr}(\boldsymbol{U}_1\boldsymbol{s}_{se,k}^\dagger\boldsymbol{s}_{se,k}^T)}{\mathrm{tr}(\overline{\boldsymbol{S}}\boldsymbol{h}_{re,k}^\dagger\boldsymbol{h}_{re,k}^T)+\mathrm{tr}((\sigma_{n_a}^2\boldsymbol{U}_1+\sigma_{n_c}^2\boldsymbol{U}_2)diag(\boldsymbol{c}_0\circ\|\boldsymbol{h}_{re,k}\|^{.2}))+\sigma_{n_{e,k}}^2} \le \tfrac{1}{\tau} \tag{5.42}
$$

Recalling the same procedure taken to deal with (P1.1-SDR), next, we devise

[2]Note that we denote problem (P2) ((P2.1),(P2.2)) with fixed $\boldsymbol{S}$ as (P2′) ((P2′.1),(P2′.2)) in the sequel.

Charnes-Cooper transformation to further convert (P2′.1-SDR) into a convex problem, denoted by (P2′.1-SDP), by replacing $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$ with $\hat{\boldsymbol{U}}_1/\xi$ and $\hat{\boldsymbol{U}}_2/\xi$, respectively. The solution for (P2′.1-SDR) is proved to be tight on the account of the following proposition.

Proposition 5.5.2. *1) The optimal solution to problem* (P2′.1-SDP) *satisfies* $\text{rank}(\hat{\boldsymbol{U}}_1^*) = 1$ *;*

2) $\hat{\boldsymbol{U}}_1^* = \hat{\boldsymbol{u}}_1^* \hat{\boldsymbol{u}}_1^{*H}$*, in which* $\hat{\boldsymbol{u}}_1^*$ *is given by*

$$\hat{\boldsymbol{u}}_1^* = \sqrt{\frac{\tau - \xi^* \sigma_{n_d}^2 - \sigma_{n_c}^2 \text{tr}(\hat{\boldsymbol{U}}_2^* \boldsymbol{C}_{rd}) - \xi^* \text{tr}(\overline{\boldsymbol{S}} \boldsymbol{h}_{rd}^\dagger \boldsymbol{h}_{rd}^T)}{\sigma_{n_a}^2 \text{tr}(\hat{\boldsymbol{u}}_1 \hat{\boldsymbol{u}}_1^H \boldsymbol{C}_{rd})}} \hat{\boldsymbol{u}}_1, \tag{5.43}$$

where $\boldsymbol{C}_{rd} = diag(\boldsymbol{c}_0 \circ \|\boldsymbol{h}_{rd}\|.^2)$*,* $\hat{\boldsymbol{u}}_1$ *and* $\boldsymbol{\Xi}'$ *are given in Appendix I;*

3) $\hat{\boldsymbol{U}}_2^*$*, of which the diagonal entries compose a vector denoted by* $\hat{\boldsymbol{u}}_2^*$*, can be modified by* $\hat{\boldsymbol{u}}_2^* \hat{\boldsymbol{u}}_2^{*H}$*.*

Proof. See Appendix I. □

The α_i^*'s are thus attained according to (5.26) via EVD of $\boldsymbol{U}_1^*$ and $\boldsymbol{U}_2^*$.

Finally, we summarize the proposed algorithm for solving problem (P2) in Table 5.1.

5.6 Secure Multi-AF Relaying: A Distributed Implementation

In this section, we study heuristic algorithms to solve (P1) and (P2) in a completely distributed fashion. Note that different from the paradigm of *distributed optimization* that allows for certain amount of information exchange based on which iterative algorithms are developed to gradually improve the system performance, we herein assume that each individual relay can only make decisions based on its local CSIs, namely, h_{sr_i}, h_{r_id}, h_{r_ie}, $\forall i$, and there is no extra means of information

Table 5.1: Algorithm for Solving (P2)

Require: $\boldsymbol{S}^*$; r^*_{SPS} that denotes the optimum value for (P1) given $\bar{\alpha}_i = .5$, $\forall i$

$ii \leftarrow 1$, $r^{(ii)}_{\text{sec}} \leftarrow r^*_{\text{SPS}}$, $r^{(0)}_{\text{sec}} \leftarrow 0$
repeat
  $ii \leftarrow ii + 1$
  $\bar{\boldsymbol{S}} \leftarrow \boldsymbol{S}^*$, $\bar{z}_i \leftarrow \text{tr}(\boldsymbol{S}^* \boldsymbol{E}_i)$, $\forall i$, and solve (P2′):
  $kk \leftarrow 0$, $r^{(0)}_{\text{DPS}} \leftarrow 10^{-6}$, $r^{(1)}_{\text{DPS}} \leftarrow 10$, $l \leftarrow \tau_{\min,2}$, $u \leftarrow 1$
  while $|r^{(kk+1)}_{\text{DPS}} - r^{(kk)}_{\text{DPS}}| \,/\, r^{(kk)}_{\text{DPS}} > \epsilon_r$ **do**
    $kk \leftarrow kk + 1$, $\tau \leftarrow \frac{l+u}{2}$
    solve (P2′.1) and
    return $H_2(\tau)$
    $r^{(kk+1)}_{\text{DPS}} \leftarrow \frac{1}{2}\log_2(\tau + H_2(\tau))$
    $r_{\text{temp}} \leftarrow \frac{1}{2}\log_2(\tilde{\tau} + H_2(\tilde{\tau}))$, where $\tilde{\tau} \leftarrow \max(\tau - \Delta\tau, \tau_{\min,2})$ and $\Delta\tau > 0$ denotes an arbitrary small value.
    if $r^{(kk+1)}_{\text{DPS}} \leq r_{\text{temp}}$ **then**
      $u \leftarrow \tau$
    else
      $l \leftarrow \tau$
    end if
  end while
  return $\boldsymbol{U}^*_1$, $\boldsymbol{U}^*_2$, and obtain $\{\alpha^*_i\}$ according to (5.26)
  $\bar{\alpha}_i \leftarrow \alpha^*_i$, $\forall i$, and solve (P1):
  $kk \leftarrow 0$, $r^{(0)}_{\text{SPS}} \leftarrow 10^{-6}$, $r^{(1)}_{\text{SPS}} \leftarrow 10$, $l \leftarrow= \tau_{\min,1}$, $u \leftarrow 1$
  while $|r^{(kk+1)}_{\text{SPS}} - r^{(kk)}_{\text{SPS}}| \,/\, r^{(kk)}_{\text{SPS}} > \epsilon_r$ **do**
    $kk \leftarrow kk + 1$, $\tau \leftarrow \frac{l+u}{2}$
    solve (P1.1) and
    return $H_1(\tau)$
    $r^{(kk+1)}_{\text{SPS}} \leftarrow \frac{1}{2}\log_2(\tau + H_1(\tau))$
    $r_{\text{temp}} \leftarrow \frac{1}{2}\log_2(\tilde{\tau} + H_1(\tilde{\tau}))$, where $\tilde{\tau} \leftarrow \max(\tau - \Delta\tau, \tau_{\min,1})$.
    if $r^{(kk+1)}_{\text{SPS}} \leq r_{\text{temp}}$ **then**
      $u \leftarrow \tau$
    else
      $l \leftarrow \tau$
    end if
  end while
  return $\boldsymbol{X}^*_1$, $\boldsymbol{S}^*$, and obtain $\{\rho^*_i\}$ and $\{\measuredangle\beta^*_i\}$ according to $w^*_{1,i} = \sqrt{1-\rho^*_i}e^{j\measuredangle\beta^*_i}$, $\forall i$
  Update $r^{(ii)}_{\text{sec}}$ according to (5.13)
until $r^{(ii)}_{\text{sec}} - r^{(ii-1)}_{\text{sec}} \leq \epsilon_0$
Ensure: $\{\alpha^*_i\}$, $\{\rho^*_i\}$, $\{\measuredangle\beta^*_i\}$, and $\boldsymbol{S}^*$

acquisition due to loads of system overhead otherwise. The reason for designing such an algorithm is as follows. On one hand, we aim for answering the question that in the least favourable situation, namely, no cooperation, what can be done to improve the achievable secrecy rate of the whole system? On the other hand, it provides lower-bound for the centralized schemes proposed in Section 5.5, which sheds light upon the trade-off between performance and complexity.

Besides, we emphasize the jamming scheme that is different from the CJ in the centralized schemes. Unlike the CJ signal that is coordinately transmitted by all relays, in the distributed implementation, each relay is only able to generated AN independently, i.e., $\boldsymbol{x}_{r2} = [\sqrt{\sigma_1}s'_1, \ldots, \sqrt{\sigma_N}s'_N]^T$, in which s_i's are $i.i.d$ AN denoted by $s'_i \sim \mathcal{CN}(0,1)$. This type of CJ is known to be IJ as mentioned in Section 5.3 with the covariance matrix $\boldsymbol{S} = diag([\sigma_1, \ldots, \sigma_N])$. In this section, we assume that each relay consumes all of its remaining power from AF for AN, i.e., $\sigma_i = \eta\rho_i\alpha_i P_s|h_{sr_i}|^2$, $\forall i \in \mathcal{N}$ (c.f. (5.5)). Hence, the AN design solely depends on α_i's and/or ρ_i's.

5.6.1 Distributed Algorithm for SPS

First, we propose a heuristic scheme for the ith AF relay to decide on ρ_i, $\forall i$, which is given by

$$\rho_i = \delta \left(1 - \frac{|h_{r_i d}|^2}{\max\limits_{k\in\mathcal{K}} |h_{r_i e,k}|^2}\right)^+, \tag{5.44}$$

where $\delta \in (0,1)$ is a constant that controls the relay's conservative level for jamming in the distributed scheme. For example, a larger δ indicates that each relay prefers to splitting a larger portion of power for jamming and vice versa. The intuition behind (5.44) is easy to see. For the ith relay, if it observes that $|h_{r_i d}|^2 \geq \max\limits_{k\in\mathcal{K}} |h_{r_i e,k}|^2$, which means that a nonnegative secrecy rate is achievable even if there is only itself in the system, it will shut down AN transmission; otherwise, it will split up to δ portion of the harvested power for jamming. For example, in an extreme case of $|h_{r_i d}|^2 \ll \max\limits_{k} |h_{r_i e,k}|^2$, probably when an eavesdropper is located within the very

proximity of this relay, it allocates δ of power to generate AN.

Next, since an individual relay cannot evaluate the secrecy performance of the whole system, $\measuredangle\beta_i$'s are simply chosen to be the optimum for the multi-AF relaying without security concerns, i.e., $\measuredangle\beta_i = -\measuredangle h_{r_i d} - \measuredangle h_{sr_i}$, $\forall i$.

5.6.2 Distributed Algorithm for DPS

Following the same designs for ρ_i's and $\measuredangle\beta_i$'s in Section 5.6.1, the remaining task for WEH-enabled relays operating with DPS is to set proper α_i's. We choose α_i's that maximize the "hypothetical SINR", where the "hypothetical" indicates that this SINR may not be the actual SINR for the destination, but just a criterion calculated based on the "hypothetical" received signal seen by the ith relay, denoted by $\tilde{y}_{d_i}$, $\forall i$, which is given by

$$\tilde{y}_{d_i} = h_{r_i d}\beta_i\sqrt{1-\alpha_i}\sqrt{P_s}h_{sr_i}s + h_{r_i d}\beta_i\sqrt{1-\alpha_i}n_{a,i} + h_{r_i d}\beta_i n_{c,i} + h_{r_i d}\sqrt{\sigma_i}s_i' + n_d. \tag{5.45}$$

The corresponding SINR is thus expressed as

$$\mathrm{SINR}_{\tilde{y}_{d_i}} = \frac{\eta(1-\rho_i)P_s|h_{sr_i}|^2}{\eta\sigma_{n_a}^2 + \frac{\eta\sigma_{n_c}^2}{1-\alpha_i} + \frac{\gamma_i(P_s|h_{sr_i}|^2+\sigma_{n_a}^2)}{\alpha_i} + \frac{\gamma_i\sigma_{n_c}^2}{\alpha_i(1-\alpha_i)} + \eta\rho_i P_s|h_{sr_i}|^2}, \ \forall i, \tag{5.46}$$

where $\gamma_i = \frac{\sigma_{n_d}^2}{P_s|h_{sr_i}|^2|h_{r_i d}|^2}$. Consequently, the maximization of (5.46) w.r.t. α_i, $\forall i$, is formulated as the following problem.

(P2-distributed) :

$$\begin{cases} \underset{\alpha_i}{\texttt{Minimize}} & \dfrac{\eta\sigma_{n_c}^2}{1-\alpha_i} + \dfrac{\gamma_i(P_s|h_{sr_i}|^2+\sigma_{n_a}^2)}{\alpha_i} + \dfrac{\gamma_i\sigma_{n_c}^2}{\alpha_i(1-\alpha_i)} \\ \texttt{Subject to} & 0 \le \alpha_i \le 1. \end{cases}$$

Proposition 5.6.1. *The optimal* α_i, $\forall i$, *to* (P2-distributed) *is given by*

$$\alpha_i^* = \frac{1}{1+\sqrt{\frac{(\eta+\gamma_i)\sigma_{nc}^2}{\gamma_i(P_s|h_{sr_i}|^2+\sigma_{na}^2+\sigma_{nc}^2)}}}. \tag{5.47}$$

Proof. It is easy to verify that problem (P2-distributed) is convex and the minimum solution of its objective function derived from the first-order derivative happens to fall within the feasible region of α_i, which is seen in (5.47). □

With ρ_i's, $\measuredangle\beta_i$'s and α_i's set, each AF relay is then able to decide its relay weight and AN transmission.

5.7 Numerical Results

In this section we compare our proposed schemes for multi-AF relaying networks operating with SPS or DPS with a variety of benchmarks. For the centralized schemes, the optimal solutions for SPS described in Section 5.5.1 is denoted by *CJ-SPS*, while Algorithm 5.1 proposed in Section 5.5.2 is denoted by *CJ-DPS*. The distributed schemes proposed in Section 5.6.1 and Section 5.6.2 are referred as *Distributed-SPS* and *Distributed-DPS*, respectively. To demonstrate the effectiveness of our AN-aided secure multi-AF relay beamforming algorithms, we also provide three benchmark schemes:*NoCJ-SPS*, *NoCJ-DPS* and *Random Power Splitting (Random PS)*. For *NoCJ-SPS*, we solve problem (P1) by replacing $\boldsymbol{S}$ with $\mathbf{0}$. Similarly, for *NoCJ-DPS*, we initialize $\overline{\boldsymbol{S}} = \mathbf{0}$ and quit the loop in Algorithm 5.1 after the very first time of solving problem (P2′). *Random PS*, on the other hand, picks up *i.i.d.* α_i's and ρ_i's uniformly generated over $[0, 1]$, respectively, and co-phases $\measuredangle\beta_i = -\measuredangle h_{sr_i} - \measuredangle h_{r_id}$, $\forall i$.

Consider that N WEH-enabled AF relays and K eavesdroppers are located within a circular area of radius R. Specifically, we assume that their respective radius and radian are drawn from uniform distributions over the interval $[0, R]$ and

$[0, 2\pi)$, respectively. we also assume that channel models consist of both large-scale path loss and small-scale multi-path fading. The unified path loss model is given by

$$L = A_0 \left(\frac{d}{d_0} \right)^{-\alpha}, \tag{5.48}$$

where $A_0 = 10^{-3}$, d denotes the relevant distance, $d_0 = 1$m is a reference distance, and α is the path loss exponent set to be 2.5. h_{sr_i}, $h_{r_i d}$, and $h_{r_i e,k}$, $\forall i \in \mathcal{N}$, $\forall k \in \mathcal{K}$, are generated from independent Rayleigh fading with zero mean and variance specified by (5.48).

The simulation parameters are set as follows unless otherwise specified: the radius defining the range is $R = 5$m; the transmit power at the source is $P_s = 40$dBm; the noise variances are set as $\sigma_{n_a}^2 = -50$dBm, $\sigma_{n_c}^2 = -80$dBm, $\sigma_{n_d}^2 = \sigma_{n_a}^2 + \sigma_{n_c}^2$, and $\sigma_{n_e,k}^2 = \sigma_{n_d}^2$, $\forall k$; the energy harvesting efficiency is assumed to be $\eta = 50\%$. In addition, the numerical examples provided below are based on an average over 500 channel realizations.

5.7.1 Secrecy Performance by Centralized Approach

In this section, we evaluate the performance of the proposed centralized designs in Section 5.5. The efficiency of the alternating optimization that iteratively attains numerical solution to problem (P2) is illustrated in Fig. 5.3, which shows the rise of the achievable secrecy rate after each round of the iteration. The most rapid increase is observed after the first iteration, which shows that the optimization of the power splitting ratios α_i's, accounts for the main factor for the secrecy rate performance gains from the SPS scheme that sets $\{\alpha_i = 0.5\}$. It is seen that the alternating algorithm converges with the relative tolerance $\epsilon_r = |r_{\text{DPS}}^{(kk+1)} - r_{\text{DPS}}^{(kk)}|/r_{\text{DPS}}^{(kk)}$ set to be 10^{-3}, after an average of 5 – 6 iterations for several channel realizations, which is reasonable in terms of iterative algorithms.

Fig. 5.4 shows the achievable secrecy rate for the legitimate Rx with the increase in the number of AF relays by different schemes. It is apparently observed that the

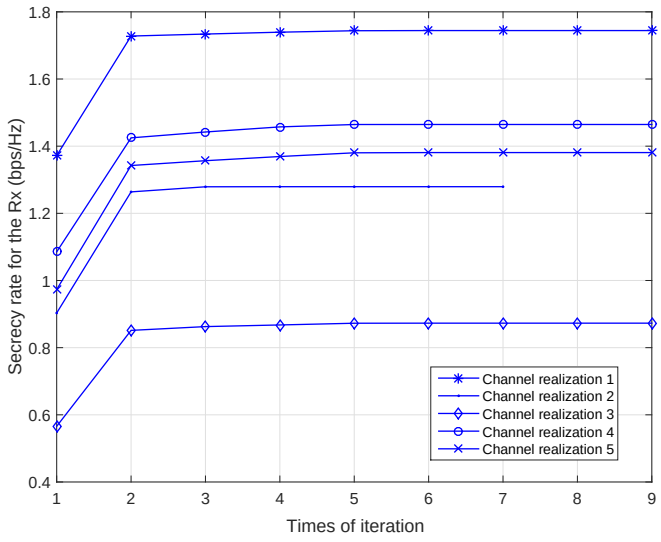
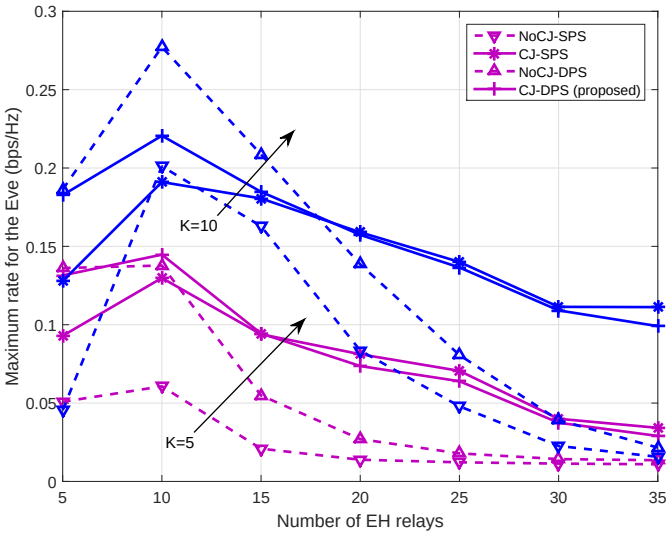


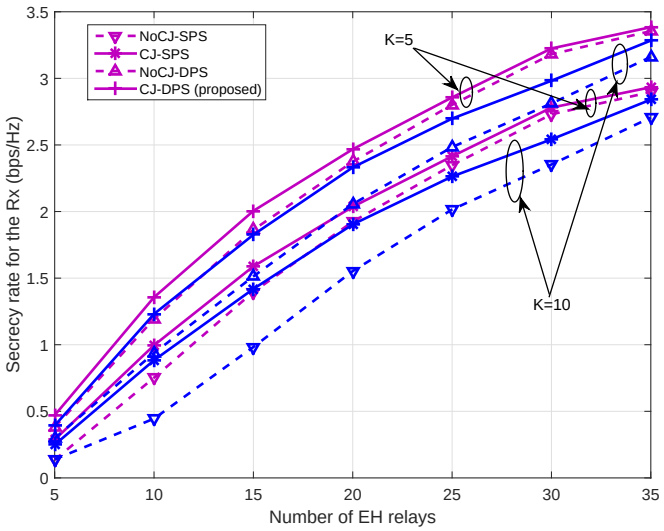
Figure 5.3: The achievable secrecy rate by *CJ-DPS* vs the number of iterations for the alternating optimization presented in Algorithm 5.1, $P_s = 40$dBm, $N = 10$, and K=5.

secure multi-AF relaying schemes assisted by the transmission of AN outperforms that w/o AN for both SPS and DPS. In addition, with the increase in N, the role of CJ gradually reduces for both schemes of SPS and DPS, which is due to the following reason. As N gets larger, the optimal designs tend to suppress the interception at the most capable eavesdropper more effectively with N d.o.f, namely, enforcing the numerator of $\text{SINR}_{\text{S,E,k}}$ and thus $r_{\text{S,E,k}}$, $\forall k \in \mathcal{K}$ to a relatively low level, which can also be observed from Fig. 5.4(a), and therefore the optimal amount of power allocated to AN beams inclines to be little otherwise the jamming yielded will be detrimental to the reception of the legitimate channel. Besides, given the same number of AF relays, the secrecy performance gains brought by the proposed schemes with CJ are more prominent in the presence of more eavesdroppers, since it is hard to reduce all the eavesdroppers' channel capacity without resort to CJ properly.

Fig. 5.5 demonstrates the achievable secrecy rate for the legitimate Rx versus the number of eavesdroppers by different schemes. First, similar to the results shown in Fig. 5.4, the proposed AN-aided multi-AF relaying designs operating with DPS-enabled relays, viz, *CJ-DPS*, perform best among all the schemes. Secondly, as



(a) $\max_{k\in\mathcal{K}} r_{\mathrm{S,E,k}}$ vs the number of AF relays.



(b) The achievable secrecy rate vs the number of AF relays.

Figure 5.4: Comparison of different schemes with $P_s = 10$dB for $K = 5$ and $K = 10$, respectively.

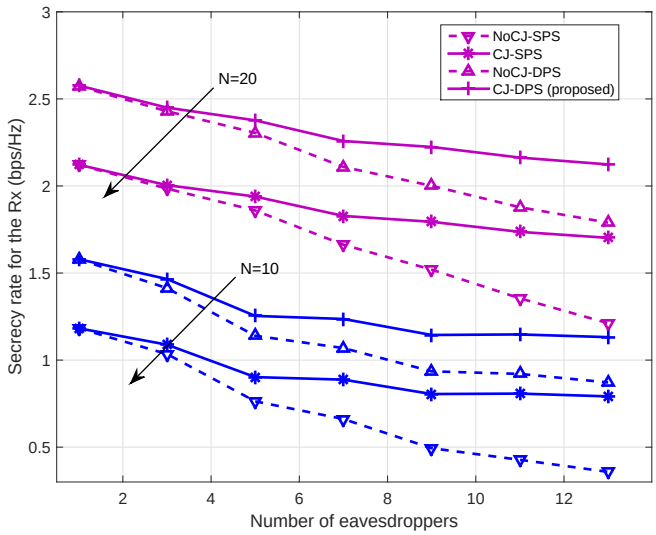


Figure 5.5: The achievable secrecy rate vs the number of eavesdroppers by different schemes with $P_s = 10$dB for $N = 10$ and $N = 20$, respectively.

K goes up, the AN-aided schemes, *CJ-DPS* and *CJ-SPS*, allow the achievable secrecy rate to drop slowly, in other words, more robust against multiple eavesdroppers, while the secrecy rate of their NoCJ counterparts almost goes down linearly with K. Moreover, with K rising, for example, more than 10, the increase in the number of relays, for example, from $N = 10$ to $N = 20$, cannot replace the role of CJ as in Fig. 5.4, since in the presence of many eavesdroppers, more relays may also result in improved eavesdroppers' decoding ability w/o the assistance of AN. It is also noteworthy that in the case of $K = 1$, there is little use of CJ by the centralized schemes, which was also observed in the numerical results of [33]. Although to the best of the authors' knowledge, there is no theoretical proof peculiar to this phenomenon in literature, a proof for no use of CJ in a special case of $N = 1$ is provided in Corollary (5.5.1).

Fig. 5.6 provides simulation results of different schemes by varying the source transmit power. It is observed that with more power available at the source, the advantage of CJ is more outstanding against the schemes w/o CJ, since given other variables fixed, larger P_s indicates larger feasible regions for problem (P1) and (P2).

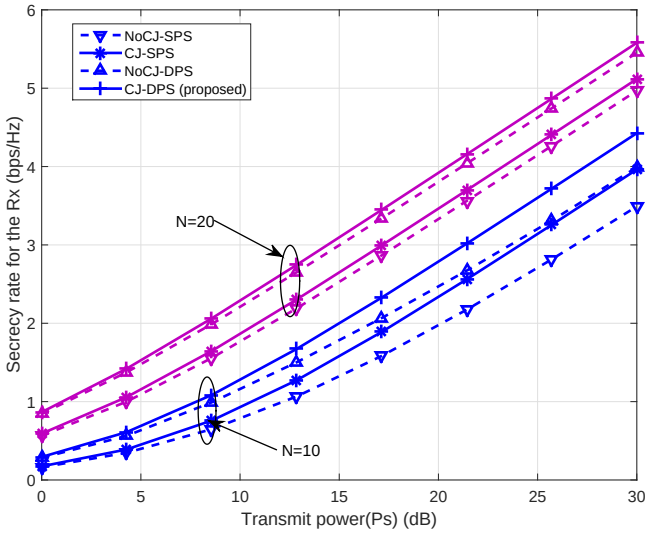


Figure 5.6: The achievable secrecy rate vs the transmit power by different schemes with K=5 for $N = 10$ and $N = 20$, respectively.

Furthermore, as similarly seen in Fig. 5.4, in a mild number of eavesdroppers (K=5), subject to the same P_s, a large number of cooperative relays enable more d.o.f in designing optimal α_i's and $\measuredangle\beta_i$'s that alleviates the dependence on AN beams to combat Eves.

5.7.2 Secrecy Performance by Distributed Implementation

In this section, we validate the purely distributed secure multi-AF relaying schemes, namely, *Distributed-SPS* and *Distributed-DPS* proposed in Section 5.6. As mentioned in Section 5.6, these heuristic designs with almost "zero" overhead incurred by information exchanged among relays are provided as benchmarks to demonstrate what can be done under the extreme "no-cooperation" circumstance, in comparison with *Random PS*. Note that any other distributed schemes with certain level of cooperation among relays are supposed to increase the secrecy performance up to the proposed centralized algorithms, namely, *CJ-DPS* and *CJ-SPS*, at the expense of extra computational complexity and system overhead. Fig. 5.7 compares the achievable secrecy rate of various schemes versus different number

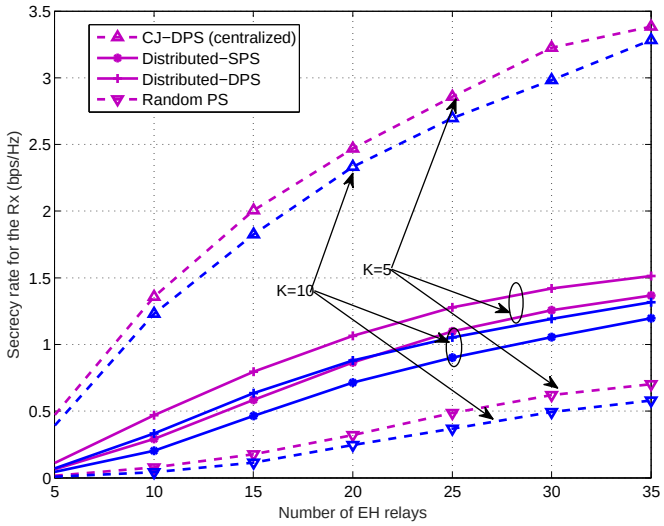


Figure 5.7: The achievable secrecy rate vs the number of AF relays by distributed algorithms with $P_s = 10$dB.

of relays. *Distributed-SPS* and *Distributed-DPS*, are seen to be outperformed by their centralized counterparts though, they are considerably superior to *Random PS*. It is also seen that the performance gap between the centralized and distributed approaches are enlarged as N increases, which is expected, since larger N yields more d.o.f for cooperation that is exclusively beneficial for the centralized schemes. Furthermore, compared with the centralized schemes, the distributed ones are more vulnerable to the increase in the eavesdroppers' number.

In Fig. 5.8, we investigate the relationship between the secrecy rate performance and the number of eavesdroppers by different methods. It can be seen that compared with the centralized schemes, the secrecy rate achieved by *Distributed-SPS* and *Distributed-DPS* both reduce more drastically with the increase in K due to the lack of effective cooperation. Also, the advantage of DPS over SPS for the distributed schemes is compromised since α_i's are not jointly designed with other parameters. At last, a similar observation has been made as that for Fig. 5.7, that is, larger N yields more visible performance gap between the centralized and distributed approaches.

In Fig. 5.9, we examine the effect of increasing the transmit power at the source

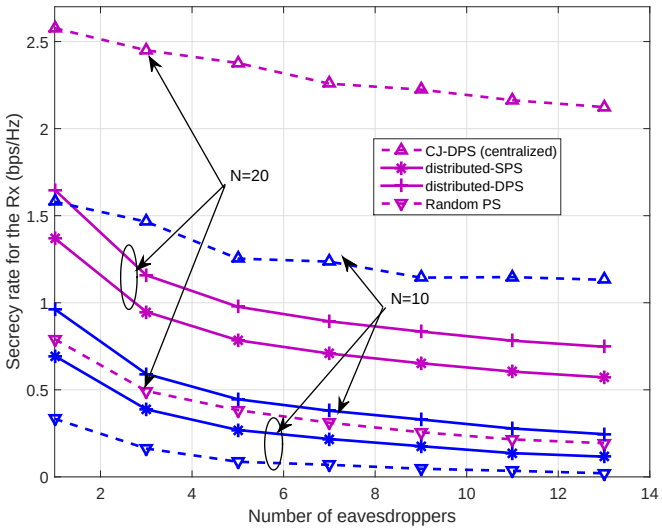


Figure 5.8: The achievable secrecy rate vs the number of eavesdroppers by distributed algorithms with $P_s = 10$dB and N=8.

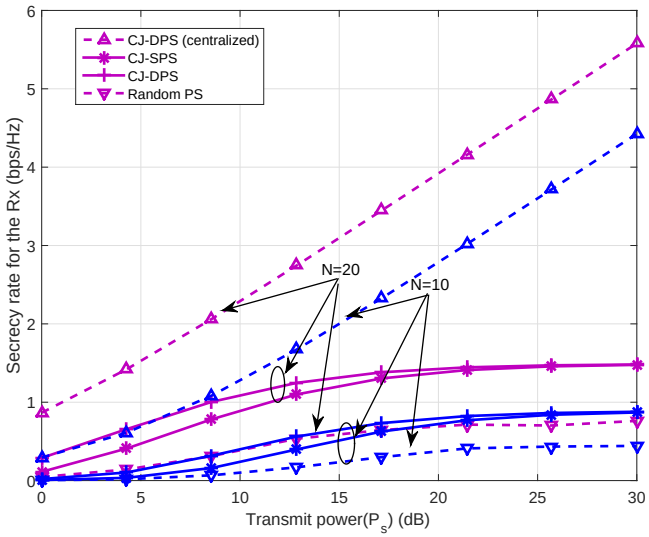


Figure 5.9: The achievable secrecy rate vs transmit power by distributed algorithms with $N = 10$ and K=5.

on the secrecy performance of different schemes under the same settings as those in Fig. 5.6. Among all the presented designs, *CJ-DPS* still achieves the best secrecy rate as observed in other examples. In addition, the fact that large N benefits more from cooperative designs is corroborated again due to the same reason as that for Fig. 5.7. Furthermore, the secrecy rate of *Distributed-SPS* or *Distributed-DPS* is quickly saturated when $P_s > 20$dB while that for their centralized counterparts still rises at a fast speed.

5.8 Secure Multi-AF Relaying: A Large Scale Realization

In this section, we consider the case when the number of EH-enabled AF relays goes to a large scale, when any centralized algorithm becomes impractical at the cost of drastically increasing computation complexity and large sum of overhead in the phase of information coordination. As an alternative, we employ in the following a simple linear relay beamforming, i.e., MF relay weights, which are completely locally decidable. Specifically, the MF coefficients are set to be $\measuredangle\beta_i = -\measuredangle h_{r_id} - \measuredangle h_{sr_i}$, $i \in \mathcal{N}$. To facilitate the analysis in the sequel, we first rewrite the channels to explicitly indicate the path loss factor and the small-scale fading as follows. $h_{sr_i} = \sqrt{\beta_{sr_i}}\bar{h}_{sr_i}$, $h_{r_id} = \sqrt{\beta_{r_id}}\bar{h}_{r_id}$, $h_{r_ie,k} = \sqrt{\beta_{r_ie,k}}\bar{h}_{r_ie,k}$, $\forall i \in \mathcal{N}$, $\forall k \in \mathcal{K}$, where β_{sr_i}, β_{r_id} and $\beta_{r_ie,k}$ are determined by (5.48), and $\bar{h}_{sr_i}$'s, $\bar{h}_{r_id}$'s and $\bar{h}_{r_ie,k}$'s are assumed to be *i.i.d.* complex Gaussian RVs with zero mean and unit variance, respectively. Next, given any α_i's, the secrecy rate analysis is developed for large N in a special case of $K = 1$ based on the following lemma.

Lemma 5.8.1. *The instantaneous achievable secrecy rate for an MF-based AF-relaying SWIPT network with a single transmission pair in the presence of one*

eavesdropper, when N goes to infinity, is given by

$$r_{\text{sec}}(\{\alpha_i\}) \xrightarrow[N\to\infty]{\text{a.s.}} \left(\frac{1}{2}\log_2\left(1+\frac{P_s|\sum_{i=1}^{N} e_i W_{-\frac{3}{4},-\frac{3}{4}}(v_i)|^2}{\sum_{i=1}^{N}\beta_{r_i d}\eta\alpha_i(\sigma_{n_a}^2+\frac{\sigma_{n_c}^2}{1-\alpha_i})(1+v_i e^{v_i}\mathrm{E_i}(-v_i))+\sigma_{n_d}^2}\right) - \frac{1}{2}\log_2\left(1+\frac{ZP_s\sum_{i=1}^{N}\beta_{r_i e,k}\beta_{sr_i}\eta\alpha_i(1-v_i-v_i^2 e^{v_i}\mathrm{E_i}(-v_i))}{\sum_{i=1}^{N}\beta_{r_i e,k}\eta\alpha_i(\sigma_{n_a}^2+\frac{\sigma_{n_c}^2}{1-\alpha_i})(1+v_i e^{v_i}\mathrm{E_i}(-v_i))+\sigma_{n_{e,k}}^2}\right)\right)^+, \tag{5.49}$$

where $v_i = \frac{(1-\alpha_i)\sigma_{n_a}^2+\sigma_{n_c}^2}{(1-\alpha_i)P_s\beta_{sr_i}}$, $e_i = \frac{\sqrt{\pi}}{2}\sqrt{\beta_{sr_i}}\sqrt{\beta_{r_i d}}\sqrt{\eta\alpha_i}v_i^{\frac{1}{4}}e^{\frac{v_i}{2}}$, $i = 1,\ldots,N$, $K = 1$, *and* Z *follows exponential distribution with unit mean. In addition,* $\mathrm{E_i}(x) = \int_{-\infty}^{x}\frac{e^t}{t}\mathrm{d}t$, $x < 0$, *stands for the exponential integral function specified by parameter* x, *and* $W_{\lambda,\mu}(\cdot)$ *represents the Whittaker function with parameters* λ *and* μ *[111, 9.22].*

Proof. See Appendix J. □

Based on Lemma 5.8.1, it follows immediately the following proposition.

Proposition 5.8.1. *The average achievable secrecy rate for an MF AF-relaying SWIPT network with a single transmission pair in the presence of one eavesdropper, when N goes to infinity, is given by*

$$\mathbb{E}[r_{\text{sec}}(\{\alpha_i\})] \xrightarrow[N\to\infty]{\text{a.s.}} \left(\frac{1}{2}\log_2(1+C_N)+\frac{1}{2\ln 2}e^{\frac{1}{D_N}}\mathrm{E_i}(-\frac{1}{D_N})\right)^+ \tag{5.50}$$

where $C_N = \frac{P_s|\sum_{i=1}^{N} e_i W_{-\frac{3}{4},-\frac{3}{4}}(v_i)|^2}{\sum_{i=1}^{N}\beta_{r_i d}\eta\alpha_i(\sigma_{n_a}^2+\frac{\sigma_{n_c}^2}{1-\alpha_i})(1+v_i e^{v_i}\mathrm{E_i}(-v_i))+\sigma_{n_d}^2}$ *and* $D_N = \frac{P_s\sum_{i=1}^{N}\beta_{r_i e,k}\beta_{sr_i}\eta\alpha_i(1-v_i-v_i^2 e^{v_i}\mathrm{E_i}(-v_i))}{\sum_{i=1}^{N}\beta_{r_i e,k}\eta\alpha_i(\sigma_{n_a}^2+\frac{\sigma_{n_c}^2}{1-\alpha_i})(1+v_i e^{v_i}\mathrm{E_i}(-v_i))+\sigma_{n_{e,k}}^2}$ *(c.f. (5.49)), $K = 1$.*

Proof. Since the instantaneous secrecy rate is non-negative if and only if $Z \le \frac{C_N}{D_N}$, (5.51) is derived by getting the expectation of (5.49) w.r.t. Z as

$$\begin{aligned}\mathbb{E}[r_{\text{sec}}(\{\alpha_i\})] &= \int_0^{\frac{C_N}{D_N}} e^{-z}\left(\frac{1}{2}\log_2(1+C_N)-\frac{1}{2}\log_2(1+ZD_N)\right)\mathrm{d}z + \int_{\frac{C_N}{D_N}}^{\infty} e^{-z}\cdot 0\,\mathrm{d}z \\ &= \frac{1}{2}\log_2(1+C_N)+\frac{1}{2\ln 2}e^{\frac{1}{D_N}}\left(\mathrm{E_i}\left(-\frac{1}{D_N}\right)-\mathrm{E_i}\left(-\frac{C_N}{D_N}\right)\right).\end{aligned} \tag{5.51}$$

Furthermore, as $C_N = \mathcal{O}(N)$ and $D_N = \mathcal{O}(1)$, $\frac{C_N}{D_N}$ approximates to $\mathcal{O}(N)$, which tends to be infinite as N goes larger, i.e., $\frac{C_N}{D_N} \xrightarrow{N\to\infty} \infty$. Accordingly, (5.51) can be further simplified as (5.50) by noting that $e^{-\frac{C_N}{D_N}} \ln(1+C_N) \xrightarrow{N\to\infty} 0$. □

Corollary 5.8.1. *In the high SNR regime in terms of* $\frac{P_s}{\sigma_{n_a}^2}$ *,* $E[r_{\text{sec}}(\{\alpha_i\})]$ *reduces to the following expression.*

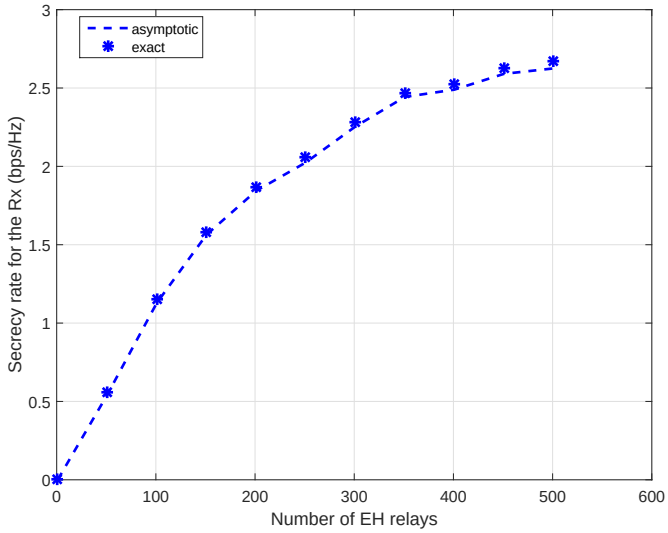
$$\mathbb{E}[r_{\text{sec}}(\{\alpha_i\})] \xrightarrow[N\to\infty]{\text{a.s.}} \frac{1}{2}\log_2(1+C'_N) + \frac{1}{2\ln 2} e^{\frac{1}{D'_N}} \text{E}_\text{i}(-\frac{1}{D'_N}), \tag{5.52}$$

where $v'_i = \frac{\sigma_{n_c}^2}{(1-\alpha_i)P_s\beta_{sr_i}}$ *(c.f.* v_i*),* $C'_N = \frac{P_s|\sum_{i=1}^N e_i W_{-\frac{3}{4},-\frac{3}{4}}(v'_i)|^2}{\sum_{i=1}^N \beta_{r_i d}\eta\alpha_i \frac{\sigma_{n_c}^2}{1-\alpha_i}(1+v'_i e^{v'_i}\text{E}_\text{i}(-v'_i))+\sigma_{n_d}^2}$*,* $\tilde{D}'_N = \frac{P_s\sum_{i=1}^N \beta_{r_i e,k}\beta_{sr_i}\eta\alpha_i(1-v'_i-v_i'^2 e^{v'_i}\text{E}_\text{i}(-v'_i))}{\sum_{i=1}^N \beta_{r_i e,k}\eta\alpha_i \frac{\sigma_{n_c}^2}{1-\alpha_i}(1+v'_i e^{v'_i}\text{E}_\text{i}(-v'_i))+\sigma_{n_{e,k}}^2}$*, and* $K = 1$.

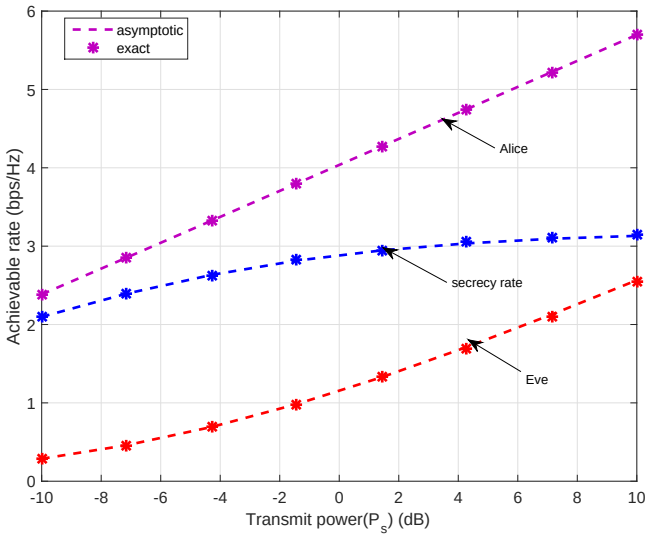
Proof. The proof can be directly obtained by removing P_s from nominator and denominator for each of the fractional terms in (5.50) and applying $\frac{P_s}{\sigma_{n_a}^2} \to \infty$. □

Under the same setup as that for a finite number of relays in Section 5.7, the asymptotic results derived above are verified by simulations in the following. Note that since N AF relays are assumed to be randomly distributed with their individual radius and radian uniformly drawn from $[0, R]$ and $[0, 2\pi)$, respectively, evaluation of the average achievable secrecy rate versus N is the mean value of 100 trials in terms of the path lass specified in (5.48), with each trial of the path loss averaged by 10^4 sub-trials in terms of the small-scale fading.

Fig. 5.10(a) shows the increasing in the average achievable secrecy rate with N. It is seen from Fig. 5.10(a) that a few hundreds of relays has already guaranteed negligible gap between the asymptotic results and the exact ones. To further investigate the performance of the MF relay beamforming by varying P_s with fixed N, Fig. 5.10(b) depicts the average achievable rate for the legitimate Rx and the eavesdropper alone, respectively, and the average achievable secrecy rate for the legitimate Rx. As seen in Fig. 5.10(a), with the simple MF precoder, the achievable rate of the Rx's channel approximately increases with P_s linearly, whereas that of



(a) The average achievable secrecy rate vs the number of AF relays.



(b) The average achievable secrecy rate vs the transmit power.

Figure 5.10: Comparison of asymptotic analysis and simulation results for $K = 1$ with $P_s = 30$dBm and $N = 200$, respectively.

the eavesdropper only admits a rough log-increase, which accounts for the growing, albeit slowly, achievable secrecy rate of the wiretap channel.

5.9 Chapter Summary

In this chapter, secure multiple single-antenna AF relaying assisted by AN transmission via CJ is studied for a SWIPT-enabled relay network, in the presence of multiple single-antenna eavesdroppers. Considering a heterogeneous power splitting (PS) protocol employed at the relay, the achievable secrecy rate for the relay wiretap channel has been maximized by jointly optimizing the complex relay beamforming weights and CJ covariance along with the PS ratios for relays operating with SPS and DPS, respectively. Reformulating the constraints into restricted hyperbolic forms essentially enabled us to develop convex optimization-based solutions. Furthermore, we also proposed an information-exchange-free distributed algorithm with very low complexity that outperforms random decisions. In addition, as a preliminary exploration for the impact of simple linear relay beamforming on the large number of relay-based average secrecy rate, the corresponding performance analysis provided guidelines for practical large-scale distributed relay implementation.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

The notion of *physical-layer security (PLS)* has advanced a paradigm shift in design of wireless security schemes to achieve information-theoretic security. The basic idea of it is to leverage the physical-layer channel induced impairments to enlarge the decoding capacity between the legitimate channel and the eavesdropper's channel so as to send confidential information to the desired receiver reliably and securely. In this thesis, an in-depth study on PLS enhancements in the flourish of self-sustaining *wireless powered communication network (WPCN)* has been given. With an emphasis on the fundamental challenges and opportunities of PLS gained by WPCN, judicious transmission and cooperation strategies along with their resource allocations have been proposed by assorted optimization technologies.

Chapter 3 is concerned with one of the fundamental challenges in *simultaneous wireless information and power transfer (SWIPT)* with security concerns, that is, the ER that is presumed to only scavenge energy from the received signal may attempt to decode the information for the IR. Under a simplified three-node SISO fading wiretap channel, with the dual usage of AN, the secrecy outage probability and ergodic secrecy capacity have been minimized and maximized for delay-limited and no-delay-limited secrecy transmission, respectively, subject to combined average and peak power constraints at the Tx, as well as an average EH constraint at the ER. For each of the two problems, a dual decomposition based optimal method was first proposed followed by an effective suboptimal algorithm, and then compared with benchmark schemes where no AN was employed or the AN was used but cannot be

canceled by the IR.

Motivated by WPCN with separately located ERs and potential eavesdroppers, the goal of wireless power transfer and achieving secrecy information transmission is not necessarily antagonistic. In the evolving densely deployed wireless networks with users's concerns of limited battery mitigated by WPCN, the advantage of cooperative strategies for improving PLS has been fundamentally boosted. Novel cooperative strategies for PLS enhancements were thus developed in Chapter 4 and 5.

A novel *harvest-and-jam (HJ)* relaying protocol is proposed in Chapter 4 to improve the secret communications over a multi-antenna AF relay wiretap channel. The secrecy rate and/or worst-case robust secrecy rate maximization problems have been formulated to jointly optimize the CJ covariance matrices of HJ helpers and beamforming matrix of the AF relay. Particularly, on imperfect CSI occasions, a new approach that equivalently models the error-bounded channel imperfections from simultaneous K HJ helpers has been proposed, for the first time, to facilitate LMI-based conic programming. Moreover, suboptimal rank-one reconstruction algorithm based on the SDR solution was proposed to strike satisfactory trade-off between complexity and performance under imperfect CSI's case.

Finally, Chapter 5 is built upon the wireless-powered CJ schemes proposed in Chapter 4 and further generalized the scope of applications to a WEH-enabled multi-AF relay network in the presence of multiple eavesdroppers, where the cooperative d.o.f was fully exploited by advocating simultaneously wireless-powered CJ and CB based on a novel hybrid PS scheme. The joint optimization of CJ and CB as well as the PS ratios was then studied with explicit relay beamforming expressions to maximize the secrecy rate for SPS and DPS-operated relays, respectively. As a by-product, in the centralized scheme with global CSI, the solution obtained for AF relay beam and the corresponding PS ratios without employing CJ was proved to achieve the global optimum by investigating the tightness of SDR.

6.2 Future Work

Several interesting future directions are highlighted as follows, which are deemed to be worthy of further investigation based on the results attained in this thesis.

The dual usage of AN for both obfuscating the potential eavesdropping ERs and satisfying their prescribed amount of WPT calls for more general channel assumptions. In particular, robust optimization framework that jointly incorporates transmit beamforming and spatially AN designs is worth investigating against imperfect CSI from both IRs and ERs. Although some work has begun studying the robust secrecy rate optimization for SWIPT-enabled MISO downlink system in the presence of one IR and multiple ERs, their channel uncertainty models are mostly limited to error-bounded deterministic ones. It is thus necessary to consider more practical channel uncertainties including random distribution-based CSI error, especially when the imperfect estimation takes main effect.

The efficiency of the proposed HJ scheme in Chapter 4 closely depends on the amount of energy harvested in the first transmission phase. Hence, instead of a two-equal-phase HJ protocol, the optimal proportion of the EH phase to the whole transmission duration is expected to be jointly optimized with other parameters. Intuitively, if the relay's main channel is already much stronger than that of the eavesdropper, very little time is likely to be allocated for EH, since the secrecy transmission achieved by the relay beamforming alone is good enough;however, if the main channel is more degraded than the eavesdropper's channel, a longer EH phase is nessary to ensure effective jamming in the following phase.

To further motivate the potential WEH-enabled helpers to assist in secret communications, practical energy and communications cooperation mechanisms need to be designed for multi-user secrecy SWIPT by game theoretic approaches. For example, ERs in the immediate vicinity of the AP could be self-interested and prefer to storing the harvested energy to providing cooperation with the Tx. In this situation, the Tx needs to offer some extra incentive, such as spectrum access in cognitive radio networks, to enable the secondary users' cooperative secrecy

transmission after they harvest energy from the AP.

Since WPCN will be beneficial for green communications beyong 5G, the advantage of key-enabling technologies in 5G, for example, massive MIMO, ought to be further studied to accommodate secrecy SWIPT. In particular, multiuser massive MIMO BC wiretap channel, in which each receiver is regarded as an eavesdropper for all messages unintended for it, has not yet been addressed. It is expected that simple linear precoding schemes may be viable in keeping the confidentiality of the desired message from all the rest of users, since they are asymptotically immune against multi-user interference in massive MIMO systems. Typically, for a two-user MIMO BC wiretap channel, the optimality of corresponding linear precoding schemes can be further evaluated by comparing it with the optimal secret dirty-paper coding (S-DPC), the secrecy capacity region of which under the matrix power constraint has been fully characterized in the literature.

Appendix A

Proof of Proposition 3.5.1

We prove Proposition 3.5.1 for the two cases of $p_1(\tilde{\alpha}) > P_{\text{peak}}$ and $p_1(\tilde{\alpha}) \leq P_{\text{peak}}$, respectively, shown as follows.

1) **Case I:** $p_1(\tilde{\alpha}) > P_{\text{peak}}$

In this case, since the minimum power for achieving r_0 already exceeds P_{peak}, the outage is inevitable. Hence,

$$L_1(p, \alpha) = 1 + (\lambda - \zeta\mu g)p. \tag{A.1}$$

To minimize $L_1(p, \alpha)$, we have

$$p^* = \begin{cases} P_{\text{peak}} & \text{if } \lambda - \zeta\mu g < 0 \\ 0 & \text{otherwise.} \end{cases} \tag{A.2}$$

Note that since in this case $X \equiv 1$, α can take any value over the interval $[0, 1]$ and thus we set $\alpha^* = 0$ for convenience.

2) **Case II:** $p_1(\tilde{\alpha}) \leq P_{\text{peak}}$

In this case, the outage can be avoided by jointly optimizing p and α. As a result, we have

$$L_1(p, \alpha) = \begin{cases} 1 + (\lambda - \zeta\mu g)p & \text{if } 0 \leq p < p_1(\tilde{\alpha}), \\ (\lambda - \zeta\mu g)p & \text{if } p_1(\tilde{\alpha}) \leq p \leq P_{\text{peak.}} \end{cases} \tag{A.3}$$

According to (A.3), the optimal power allocation to minimize $L_1(p, \alpha)$ also depends on whether $\lambda - \zeta\mu g < 0$ or not. Thus, in the following we further

discuss two subcases.

- **Subcase II-1:** $\lambda - \zeta\mu g < 0$. In this subcase, given any $\alpha = \bar{\alpha}$ with $p_1(\bar{\alpha}) \leq P_{\text{peak}}$, $L_1(p, \bar{\alpha})$ is a monotonically decreasing function over p. As a result, over the interval $0 \leq p \leq p_1(\bar{\alpha})$, $L_1(p, \bar{\alpha})$ is minimized by $p = p_1(\bar{\alpha})$; while over the interval $p_1(\bar{\alpha}) < p \leq P_{\text{peak}}$, it is minimized by $p = P_{\text{peak}}$. Note that given any $\bar{\alpha}$ with $p_1(\bar{\alpha}) \leq P_{\text{peak}}$, it follows that $1 + (\lambda - \zeta\mu g)p_1(\bar{\alpha}) > (\lambda - \zeta\mu g)P_{\text{peak}}$. Therefore, the optimal power allocation for any $\bar{\alpha}$ is $p^* = P_{\text{peak}}$. Moreover, any $\bar{\alpha}$ that satisfies $p_1(\bar{\alpha}) \leq P_{\text{peak}}$ is optimal.

- **Subcase II-2:** $\lambda - \zeta\mu g \geq 0$. In this subcase, given any $\alpha = \bar{\alpha}$ with $p_1(\bar{\alpha}) \leq P_{\text{peak}}$, $L_1(p, \bar{\alpha})$ is a monotonically increasing function over p. As a result, over the interval $0 \leq p < p_1(\bar{\alpha})$, $L_1(p, \bar{\alpha})$ is minimized by $p = 0$ (i.e., $L_1^*(p, \bar{\alpha}) = 1$); while over the interval $p_1(\bar{\alpha}) \leq p \leq P_{\text{peak}}$, it is minimized by $p = p_1(\bar{\alpha})$. Furthermore, $p_1(\bar{\alpha})$ can be minimized by setting $\bar{\alpha} = \tilde{\alpha}$ (i.e., $L_1^*(p, \bar{\alpha}) = (\lambda - \zeta\mu g)p_1(\tilde{\alpha})$). Hence, the optimal power allocation for minimizing $L_1(p, \bar{\alpha})$ depends on the relationship between 1 and $(\lambda - \zeta\mu g)p_1(\tilde{\alpha})$. If $1 < (\lambda - \zeta\mu g)p_1(\tilde{\alpha})$, since $p^* = 0$, any $\bar{\alpha}$ is optimal and thus we set $\alpha^* = 0$ for simplicity; however, if $1 \geq (\lambda - \zeta\mu g)p_1(\tilde{\alpha})$, the optimal power allocation is $p^* = p_1(\tilde{\alpha})$ with the optimal power splitting ratio $\alpha^* = \tilde{\alpha}$.

By combing the above two cases of $p_1(\tilde{\alpha}) > P_{\text{peak}}$ and $p_1(\tilde{\alpha}) \leq P_{\text{peak}}$, Proposition 3.5.1 is thus proved.

Appendix B

Proof of Proposition 3.5.2

According to (3.10), the derivative of $R(\alpha, \bar{p})$ over α is given by

$$\frac{\partial R(\alpha, \bar{p})}{\partial \alpha} = \begin{cases} \frac{(1-2\alpha+x)hg\bar{p}^2}{\ln 2(\sigma_1^2+(1-\alpha)h\bar{p})(\sigma_2^2+\alpha g\bar{p})} & \text{if } \alpha \geq x, \\ 0 & \text{otherwise,} \end{cases} \tag{B.1}$$

where $x = \frac{\sigma_1^2}{h\bar{p}} - \frac{\sigma_2^2}{g\bar{p}}$. It can be shown from (B.1) that if $x < -1$, then $\frac{\partial R(\alpha,\bar{p})}{\partial \alpha} < 0$ with $0 \leq \alpha \leq 1$. Thus, $R(\alpha, \bar{p})$ is a monotonically decreasing function over α in the interval $[0, 1]$, and the optimal solution to problem (3.26) is $\alpha^* = 0$. If $-1 \leq x < 1$, it can be shown that $R(\alpha, \bar{p})$ is a non-decreasing function of α over the interval $[0, \frac{1}{2} + \frac{x}{2}]$, but a monotonically decreasing function over $(\frac{1}{2} + \frac{x}{2}, 1]$. As a result, we have $\alpha^* = \frac{1}{2} + \frac{x}{2}$. Finally, if $x \geq 1$, $\frac{\partial R(\alpha,\bar{p})}{\partial \alpha} \geq 0$, and thus $R(\alpha, \bar{p})$ is non-decreasing over $\alpha \in [0, 1]$. In this case, the optimal solution to problem (P1.2-sub) is $\alpha^* = 1$. Proposition 3.5.2 is thus proved.

Appendix C

Proof of Proposition 4.4.1

The KKT conditions of (P1′.1-RW-SDR) are given by

$$\boldsymbol{A}^*\boldsymbol{X}^* = 0, \tag{C.1a}$$

$$\boldsymbol{B}_k^*\boldsymbol{Q}_k^* = 0, \forall k, \tag{C.1b}$$

$$\beta_k^* \left(\mathrm{tr}(\boldsymbol{Q}_k^*) - \tau^*\eta P_s\|\boldsymbol{h}_k\|^2\right) = 0, \forall k. \tag{C.1c}$$

According to (4.27), if for certain k, $\beta_k^* = 0$, then $\boldsymbol{B}_k^* = -\lambda^*\tilde{\boldsymbol{h}}_k^*\tilde{\boldsymbol{h}}_k^T + \alpha^*\bar{\gamma}_e\boldsymbol{g}_k^*\boldsymbol{g}_k^T$ and thus $\mathrm{rank}(\boldsymbol{B}_k^*) \leq \mathrm{rank}(\tilde{\boldsymbol{h}}_k^*\tilde{\boldsymbol{h}}_k^T) + \mathrm{rank}(\boldsymbol{g}_k^*\boldsymbol{g}_k^T) = 2$, which yields $\mathrm{rank}(\boldsymbol{Q}_k^*) \geq N_t - 2$ as a result of (C.1b). Otherwise, when $\beta_k^* > 0$, we will have $\mathrm{rank}(\boldsymbol{B}_k^*) \geq \mathrm{rank}(-\beta_k^*\boldsymbol{I} - \lambda^*\tilde{\boldsymbol{h}}_k^*\tilde{\boldsymbol{h}}_k^T) - \mathrm{rank}(\alpha^*\bar{\gamma}_e\boldsymbol{g}_k^*\boldsymbol{g}_k^T) = N_t - 1$ [57, *Lemma A.1*], which implies $\mathrm{rank}(\boldsymbol{Q}_k^*) \leq 1$. However, $\mathrm{rank}(\boldsymbol{Q}_k^*)$ cannot be 0, since otherwise $\mathrm{tr}(\boldsymbol{Q}_k^*) - \tau^*\eta P_s\|\boldsymbol{h}_k\|^2 < 0$ and thus $\beta_k^* = 0$ according to (C.1c), which contradicts to $\beta_k^* > 0$. Hence, when $\beta_k^* > 0$, $\mathrm{rank}(\boldsymbol{Q}_k^*) = 1$.

Next, define $\boldsymbol{C}^* = -\lambda^*\sigma_r^2\overline{\boldsymbol{Y}}_1 - \alpha^*P_s\boldsymbol{F}_2 + \alpha^*\bar{\gamma}_e\sigma_r^2\overline{\boldsymbol{Y}}_2 - \beta_0^*\overline{\boldsymbol{\Phi}}$ and according to (4.26), we have

$$\boldsymbol{A}^* = P_s\boldsymbol{F}_1 + \boldsymbol{C}^*. \tag{C.2}$$

Then define r_c, $\boldsymbol{\Xi}$ and $\boldsymbol{\eta}_n$, $n = 1, \ldots, N_t^2 - r_c$ (c.f. (4.29)). Similar to the approach used in [57, Appendix B], we discuss the structure of the optimal $\boldsymbol{X}$ under two cases.

(1) **Case I**: $r_c = N_t^2$

As $\boldsymbol{C}^*$ is full-rank, $\mathrm{rank}(\boldsymbol{A}^*) \geq r_c - 1 = N_t^2 - 1$ and hence $N_t^2 - 1 \leq \mathrm{rank}(\boldsymbol{A}^*) \leq N_t^2$. If $\mathrm{rank}(\boldsymbol{A}^*) = N_t^2 - 1$, $\mathrm{rank}(\mathbf{null}(\boldsymbol{A}^*)) = 1$ and it follows that $\boldsymbol{X}^* = b\boldsymbol{\xi}\boldsymbol{\xi}^H$

by assuming $\boldsymbol{\xi}$ as the only basis of $\mathbf{null}(\boldsymbol{A}^*)$. Otherwise, according to (C.1a), we obtain $\boldsymbol{X}^* = \mathbf{0}$, which ceases the secrecy transmission and cannot be the optimal solution to (P1′.1-RW-SDR).

(2) **Case II**: $r_c < N_t^2$

If $\boldsymbol{C}^*$ is not full-rank, $\mathrm{rank}(\boldsymbol{A}^*) \geq r_c - 1$. Then by pre-multiplying $\boldsymbol{\eta}_n^H$ and post-multiplying $\boldsymbol{\eta}_n \in \boldsymbol{\Xi}$ with both sides of (C.2), we have

$$\boldsymbol{\eta}_n^H \boldsymbol{A}^* \boldsymbol{\eta}_n = P_s \boldsymbol{\eta}_n^H \boldsymbol{F}_1 \boldsymbol{\eta}_n + \boldsymbol{\eta}_n^H \boldsymbol{C}^* \boldsymbol{\eta}_n = P_s \boldsymbol{\eta}_n^H \boldsymbol{F}_1 \boldsymbol{\eta}_n, \ \forall n. \tag{C.3}$$

According to (4.25), it is necessary for $\boldsymbol{A}^* \preceq \mathbf{0}$ to obtain an optimal solution of $\boldsymbol{X}^*$ and therefore $\boldsymbol{\eta}_n^H \boldsymbol{A}^* \boldsymbol{\eta}_n \leq 0$, which conforms to $P_s \boldsymbol{\eta}_n^H \boldsymbol{F}_1 \boldsymbol{\eta}_n \geq 0$ if and only if $\boldsymbol{A}^* \boldsymbol{\eta}_n = 0$ and $\boldsymbol{F}_1 \boldsymbol{\eta}_n = 0$. Hence, $\boldsymbol{\Xi} \subseteq \mathbf{null}(\boldsymbol{A}^*)$, i.e., $N_t^2 - \mathrm{rank}(\boldsymbol{A}^*) \geq N_t^2 - r_c \Rightarrow \mathrm{rank}(\boldsymbol{A}^*) \leq r_c$. Next, we show $\mathrm{rank}(\boldsymbol{A}^*) \neq r_c$ by contradiction. If $\mathrm{rank}(\boldsymbol{A}^*) = r_c$, $\boldsymbol{\Xi} = \mathbf{null}(\boldsymbol{A}^*)$, and $\boldsymbol{X}^* = \sum_{n=1}^{N_t^2 - r_c} a_n \boldsymbol{\eta}_n \boldsymbol{\eta}_n^H$. However, in this case, since $\boldsymbol{F}_1 \boldsymbol{\eta}_n = 0$, $P_s \mathrm{tr}(\boldsymbol{F}_1 \boldsymbol{X}^*) = 0$, which is apparently not optimal. Hence, we have $\mathrm{rank}(\boldsymbol{A}^*) = r_c - 1$ and thus $\mathrm{rank}(\mathbf{null}(\boldsymbol{A}^*)) = N_t^2 - r_c + 1$. This indicates that besides the basis in $\boldsymbol{\Xi}$, $\mathbf{null}(\boldsymbol{A}^*)$ spans over an extra dimension of basis, which is denoted by $\boldsymbol{\xi}$, and hence $\boldsymbol{X}^* = \sum_{n=1}^{N_t^2 - r_c} a_n \boldsymbol{\eta}_n \boldsymbol{\eta}_n^H + b \boldsymbol{\xi} \boldsymbol{\xi}^H$.

Assume that $(\boldsymbol{X}^*, \{\boldsymbol{Q}_k^*\}, \tau^*)$ is the optimal solution to (P1′.1-RW-SDR) with $\mathrm{rank}(\boldsymbol{X}^*) > 1$. Then construct a new solution $\{\hat{\boldsymbol{X}}^*, \hat{\boldsymbol{Q}}_k^*, \hat{\tau}^*\}$ according to (4.30)–(4.32). Now, we check if the reconstructed solution is feasible if (4.33) holds.

First,

$$\begin{aligned}
&\sigma_r^2 \mathrm{tr}(\overline{\boldsymbol{Y}}_1 \hat{\boldsymbol{X}}^*) + \sum_{k=1}^{K} \tilde{\boldsymbol{h}}_k^T \hat{\boldsymbol{Q}}_k^* \tilde{\boldsymbol{h}}_k^\dagger + \hat{\tau}^* \sigma_b^2 \\
&\leq \sigma_r^2 \mathrm{tr}\left(\overline{\boldsymbol{Y}}_1 \left(\boldsymbol{X}^* - \sum_{n=1}^{N_t^2 - r_c} a_n \boldsymbol{\eta}_n \boldsymbol{\eta}_n^H\right)\right) \\
&+ \sum_{k=1}^{K} \tilde{\boldsymbol{h}}_k^T \boldsymbol{Q}_k^* \tilde{\boldsymbol{h}}_k^\dagger + \left(\tau^* + \tfrac{\sigma_r^2}{\sigma_b^2} \sum_{n=1}^{N_t^2 - r_c} a_n \mathrm{tr}(\overline{\boldsymbol{Y}}_1 \boldsymbol{\eta}_n \boldsymbol{\eta}_n^H)\right) \sigma_b^2 \\
&= \sigma_r^2 \mathrm{tr}(\overline{\boldsymbol{Y}}_1 \boldsymbol{X}^*) + \sum_{k=1}^{K} \tilde{\boldsymbol{h}}_k^T \boldsymbol{Q}_k^* \tilde{\boldsymbol{h}}_k^\dagger + \tau^* \sigma_b^2 \overset{(a)}{\leq} 1.
\end{aligned} \tag{C.4}$$

Moreover,

$$\begin{aligned}
&P_s \mathrm{tr}(\boldsymbol{F}_2 \hat{\boldsymbol{X}}^*) = P_s \mathrm{tr}\left(\boldsymbol{F}_2 (\boldsymbol{X}^* - \sum_{n=1}^{N_t^2 - r_c} a_n \boldsymbol{\eta}_n \boldsymbol{\eta}_n^H)\right) \\
&\overset{(b)}{\leq} \bar{\gamma}_e \left(\sigma_r^2 \mathrm{tr}(\overline{\boldsymbol{Y}}_2 \boldsymbol{X}^*) + \sum_{k=1}^{K} \boldsymbol{g}_k^T \boldsymbol{Q}_k^* \boldsymbol{g}_k^\dagger + \tau^* \sigma_e^2\right) + \bar{\gamma}_e \left(\sigma_e^2 \Delta\tau - \sigma_r^2 \mathrm{tr}\left(\overline{\boldsymbol{Y}}_2 \sum_{n=1}^{N_t^2 - r_c} a_n \boldsymbol{\eta}_n \boldsymbol{\eta}_n^H\right)\right) \\
&= \bar{\gamma}_e \left(\sigma_r^2 \mathrm{tr}(\overline{\boldsymbol{Y}}_2 \hat{\boldsymbol{X}}^*) + \sum_{k=1}^{K} \boldsymbol{g}_k^T \hat{\boldsymbol{Q}}_k^* \boldsymbol{g}_k^\dagger + \hat{\tau}^* \sigma_e^2\right).
\end{aligned} \tag{C.5}$$

In addition, (4.24c)–(4.24e) are easily shown to be satisfied. In the above, (a) and (b) hold due to the feasibility in (4.24a) and (4.24b), respectively. Further, $P_s \mathrm{tr}(\boldsymbol{F}_1 \hat{\boldsymbol{X}}^*) = P_s \mathrm{tr}(\boldsymbol{F}_1 \boldsymbol{X}^*)$ shows that the reconstructed solution achieves the same optimum value as that of (P1′.1-RW-SDR). Hence, an optimal solution to (P1′.1-RW-SDR) with rank-one $\boldsymbol{X}$ is ensured.

Appendix D

Proof of Proposition 4.4.2

Denoting the dual variable associated with (4.41a), (4.41b) and (4.41c) by λ, α and β_0, respectively, the Lagrangian of (P1′.1-sub2-SDR) is expressed as

$$L(\chi) = \mathrm{tr}\left((P_s\boldsymbol{F}_1 - \lambda\sigma_r^2\overline{\boldsymbol{Y}}_1 - \alpha P_s\boldsymbol{F}_2 + \alpha\bar{\gamma}_e\sigma_r^2\overline{\boldsymbol{Y}}_2 - \beta_0\overline{\boldsymbol{\Phi}})\boldsymbol{X}\right) + \left(-\lambda\sigma_b^2 + \alpha\bar{\gamma}_e(q+\sigma_e^2) + \beta_0 P_r\right)\tau + \lambda, \quad \text{(D.1)}$$

where $\chi = \{\boldsymbol{X}, \tau, \lambda, \alpha, \beta_0\}$ denotes the set consisting of all the primal and dual variables. Since problem
(P1′.1-sub2-SDR) satisfies the Slater condition, its optimum value admits zero duality gap with its dual counterpart. Furthermore, according to (D.1), in order for the dual function to be bounded from above, the following constraints must hold:

$$\boldsymbol{Z} = P_s\boldsymbol{F}_1 - \lambda\sigma_r^2\overline{\boldsymbol{Y}}_1 - \alpha P_s\boldsymbol{F}_2 + \alpha\bar{\gamma}_e\sigma_r^2\overline{\boldsymbol{Y}}_2 - \beta_0\overline{\boldsymbol{\Phi}} \preceq \boldsymbol{0}, \quad \text{(D.2)}$$

$$-\lambda\sigma_b^2 + \alpha\bar{\gamma}_e(q+\sigma_e^2) + \beta_0 P_r \leq 0. \quad \text{(D.3)}$$

The dual problem is therefore given by

$$\text{(D-P1′.1-sub2-SDR)} : \min_{\lambda,\alpha,\beta_0} \lambda$$

$$\texttt{s.t.}\ \text{(D.2)}, \text{(D.3)}, \quad \text{(D.4a)}$$

$$(\lambda, \alpha, \beta_0)^T \geq \boldsymbol{0}. \quad \text{(D.4b)}$$

It is observed that $\boldsymbol{Z}$ is of the same form as the Hessian matrix with respect to $\boldsymbol{X}$ without rank relaxation. According to [112, Theorem 2.1], $\boldsymbol{Z} \preceq \boldsymbol{0}$ implies that

the SDR problem (P1′.1-sub2-SDR) is tight in this case, i.e., $\exists \boldsymbol{w}^*$ such that $\boldsymbol{X}^* = \boldsymbol{w}^* \boldsymbol{w}^{*H}$. Moreover, since KKT condition necessitates $\boldsymbol{Z}^* \boldsymbol{X}^* = \mathbf{0}$, it follows that $\boldsymbol{w}^*$ is the eigenvector corresponds to the zero-eigenvalue of $\boldsymbol{Z}^*$. Hence, we have $\boldsymbol{w}^* = \mu \nu_{\max}(\boldsymbol{Z}^*)$, where $\mu = \sqrt{\frac{P_r}{\mathrm{tr}(\overline{\boldsymbol{\Phi}}) \boldsymbol{\nu}_{\max}(\boldsymbol{Z}^*) \boldsymbol{\nu}_{\max}^H(\boldsymbol{Z}^*)}}$ is due to the power constraint of (4.24c), which completes the proof.

Appendix E

Proof of Proposition 4.5.1

First, given $\tilde{\boldsymbol{h}}_k$, $k = 2, \ldots, K$, fixed, only consider the uncertainty of $\tilde{\boldsymbol{h}}_1$. Since $\|\Delta\tilde{\boldsymbol{h}}_1\|_2^2 \leq \epsilon_1''$, we have $1 - \frac{(\Delta\tilde{\boldsymbol{h}}_1^\dagger)^H \Delta\tilde{\boldsymbol{h}}_1^\dagger}{\epsilon_1''} \geq 0$. By applying Lemma 4.5.2 to (4.60) with $\boldsymbol{H}_1{}^{(1)} = P_s\boldsymbol{X}'' - \delta\sigma_r^2\boldsymbol{X}' + w^{(0)}\boldsymbol{I}$, $\boldsymbol{F}_1{}^{(1)} = (P_s\boldsymbol{X}'' - \delta\sigma_r^2\boldsymbol{X}')\hat{\tilde{\boldsymbol{h}}}^\dagger$, $\boldsymbol{G}_1{}^{(1)} = \boldsymbol{0}$, $c_1^{(1)} = \hat{\tilde{\boldsymbol{h}}}^T(P_s\boldsymbol{X}'' - \delta\sigma_r^2\boldsymbol{X}')\hat{\tilde{\boldsymbol{h}}}^\dagger - \delta\hat{\tilde{\boldsymbol{h}}}_1^T\boldsymbol{Q}_1\hat{\tilde{\boldsymbol{h}}}_1^\dagger - \delta\sum_{i=2}^K \tilde{\boldsymbol{h}}_i^T\boldsymbol{Q}_i\tilde{\boldsymbol{h}}_i^\dagger - \delta\sigma_b^2 - w^{(0)}N_t\epsilon_0'$, $\boldsymbol{B}_1{}^{(1)} = -\delta\boldsymbol{Q}_1\hat{\tilde{\boldsymbol{h}}}_1^\dagger$, and $\boldsymbol{A}_1{}^{(1)} = -\delta\boldsymbol{Q}_1$, there exists $w^{(1)} \geq 0$ such that the following LMI holds:

$$\begin{bmatrix} \boldsymbol{H}_1^{(1)} & \boldsymbol{F}_1{}^{(1)} & \boldsymbol{G}_1{}^{(1)} \\ \boldsymbol{F}_1{}^{(1)H} & c_1^{(1)} & \boldsymbol{B}_1{}^{(1)H} \\ \boldsymbol{G}_1{}^{(1)H} & \boldsymbol{B}_1{}^{(1)} & \boldsymbol{A}_1{}^{(1)} \end{bmatrix} - w^{(1)} \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \frac{-\boldsymbol{I}}{\epsilon_1''} \end{bmatrix} \succeq \boldsymbol{0}. \tag{E.1}$$

Note that for $\boldsymbol{Q}_1 \succeq \boldsymbol{0}$, there always exists $w^{(1)} > 0$ such that $\frac{w^{(1)}\boldsymbol{I}}{\epsilon_1''} + \boldsymbol{A}_1{}^{(1)} \succ \boldsymbol{0}$ and we assume that such constraint is applied. According to the property of Schur-Complements [84, A. 5.5], for (E.1), we have

$$\begin{cases} \begin{bmatrix} \boldsymbol{H}_1{}^{(1)} & \boldsymbol{F}_1{}^{(1)} \\ \boldsymbol{F}_1{}^{(1)H} & c_1^{(1)} - w^{(1)} \end{bmatrix} \\ - \begin{bmatrix} \boldsymbol{G}_1{}^{(1)} \\ \boldsymbol{B}_1{}^{(1)H} \end{bmatrix} \left(\boldsymbol{A}_1{}^{(1)} + \frac{w^{(1)}\boldsymbol{I}}{\epsilon_1''}\right)^{-1} \begin{bmatrix} \boldsymbol{G}_1{}^{(1)H} & \boldsymbol{B}_1{}^{(1)} \end{bmatrix} \succeq \boldsymbol{0}, \\ \frac{w^{(1)}\boldsymbol{I}}{\epsilon_1''} + \boldsymbol{A}_1{}^{(1)} \succ \boldsymbol{0}, \end{cases} \tag{E.2}$$

which can be reexpressed as

$$\begin{bmatrix} \boldsymbol{A}_1{}^{(1)} + \frac{w^{(1)}\boldsymbol{I}}{\epsilon_1''} & \boldsymbol{G}_1{}^{(1)H} & \boldsymbol{B}_1{}^{(1)} \\ \boldsymbol{G}_1{}^{(1)} & \boldsymbol{H}_1{}^{(1)} & \boldsymbol{F}_1{}^{(1)} \\ \boldsymbol{B}_1{}^{(1)H} & \boldsymbol{F}_1{}^{(1)H} & c_1^{(1)} - w^{(1)} \end{bmatrix} \succeq \boldsymbol{0}. \tag{E.3}$$

Next, assume that the robust design for (4.60) has been considered against the precedent $k-1$ uncertainties, i.e.,

$$\begin{bmatrix} \boldsymbol{H}_1{}^{(k-1)} & \boldsymbol{F}_1{}^{(k-1)} & \boldsymbol{G}_1{}^{(k-1)} \\ \boldsymbol{F}_1{}^{(k-1)H} & c_1^{(k-1)} & \boldsymbol{B}_1{}^{(k-1)H} \\ \boldsymbol{G}_1{}^{(k-1)H} & \boldsymbol{B}_1{}^{(k-1)} & \boldsymbol{A}_1{}^{(k-1)} \end{bmatrix} - w^{(k-1)} \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{-\boldsymbol{I}}{\epsilon''_{k-1}} \end{bmatrix} \succeq \mathbf{0},\ k \geq 2. \tag{E.4}$$

Applying a similar procedure as that for (E.1), (E.4) can be recast as

$$\begin{bmatrix} \frac{w^{(k-1)}\boldsymbol{I}}{\epsilon''_{k-1}} + \boldsymbol{A}_1{}^{(k-1)} & \boldsymbol{G}_1{}^{(k-1)H} & \boldsymbol{B}_1{}^{(k-1)} \\ \boldsymbol{G}_1{}^{(k-1)} & \boldsymbol{H}_1{}^{(k-1)} & \boldsymbol{F}_1{}^{(k-1)} \\ \boldsymbol{B}_1{}^{(k-1)H} & \boldsymbol{F}_1{}^{(k-1)H} & c_1^{(k-1)} - w^{(k-1)} \end{bmatrix} \succeq \mathbf{0}. \tag{E.5}$$

Then given $\tilde{\boldsymbol{h}}_i,\ i = k+1, \ldots, K$ fixed, accommodate the kth uncertainty, i.e., $\tilde{\boldsymbol{h}}_k \in \tilde{\mathcal{H}}_k$, for (E.5). By applying Lemma 4.5.2 to the uncertainty of $\tilde{\boldsymbol{h}}_k$, the implication $\|\Delta\tilde{\boldsymbol{h}}_k\|_2^2 \leq \epsilon''_k \Rightarrow$ (E.5) holds if and only if there exists $w^{(k)} \geq 0$ such that

$$\begin{bmatrix} \boldsymbol{H}_1{}^{(k)} & \boldsymbol{F}_1{}^{(k)} & \boldsymbol{G}_1{}^{(k)} \\ \boldsymbol{F}_1{}^{(k)H} & c_1^{(k)} & \boldsymbol{B}_1{}^{(k)H} \\ \boldsymbol{G}_1{}^{(k)H} & \boldsymbol{B}_1{}^{(k)} & \boldsymbol{A}_1{}^{(k)} \end{bmatrix} - w^{(k)} \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{-\boldsymbol{I}}{\epsilon''_k} \end{bmatrix} \succeq \mathbf{0}, \tag{E.6}$$

where

$$\begin{aligned} \boldsymbol{H}_1{}^{(k)} &= \begin{bmatrix} \boldsymbol{A}_1{}^{(k-1)} + \frac{w^{(k-1)}\boldsymbol{I}}{\epsilon''_{k-1}} & \boldsymbol{G}_1{}^{(k-1)H} \\ \boldsymbol{G}_1{}^{(k-1)} & \boldsymbol{H}_1{}^{(k-1)} \end{bmatrix}, \\ \boldsymbol{F}_1{}^{(k)} &= \begin{bmatrix} \boldsymbol{B}_1{}^{(k-1)} \\ \boldsymbol{F}_1{}^{(k-1)} \end{bmatrix},\ \boldsymbol{G}_1{}^{(k)} = \mathbf{0}, \end{aligned} \tag{E.7}$$

$c_1^{(k)} = \hat{\tilde{\boldsymbol{h}}}^T(P_s\boldsymbol{X}'' - \delta\sigma_r^2\boldsymbol{X}')\hat{\tilde{\boldsymbol{h}}}^\dagger - \delta\sum_{j=1}^{k}\hat{\tilde{\boldsymbol{h}}}_j^T\boldsymbol{Q}_j\hat{\tilde{\boldsymbol{h}}}_j^\dagger - \delta\sum_{i=k+1}^{K}\tilde{\boldsymbol{h}}_i^T\boldsymbol{Q}_i\tilde{\boldsymbol{h}}_i^\dagger - \delta\sigma_b^2 - w^{(0)}N_t\epsilon'_0 - \sum_{l=1}^{k-1}w^{(l)}$, $\boldsymbol{B}_1{}^{(k)} = -\delta\boldsymbol{Q}_k\hat{\tilde{\boldsymbol{h}}}_k^\dagger$ and $\boldsymbol{A}_1{}^{(k)} = -\delta\boldsymbol{Q}_k$, $k \geq 2$. Thus, using the method of mathematical induction, (4.60) holds for $\tilde{\boldsymbol{h}}_k \in \tilde{\mathcal{H}}_k$, $k = 1, \ldots, K$, if and only if there exists $\{w(k) \geq 0\}$, such that (4.61) is satisfied, which completes the proof.

Appendix F

Proof of Proposition 4.5.2

Taking the similar procedure as that for dealing with (4.59), the implication $\|\Delta \boldsymbol{g}\|^2 \leq N_t \epsilon_0 \Rightarrow$ (4.64) holds if and only if there exists $v^{(0)} \geq 0$ such that the following LMI holds:

$$\begin{bmatrix} \boldsymbol{H}_2 & \boldsymbol{F}_2 \\ \boldsymbol{F}_2^H & c_2 \end{bmatrix} \succeq \boldsymbol{0}, \tag{F.1}$$

where $\boldsymbol{H}_2 = -P_s \boldsymbol{X}'' + \bar{\gamma}_e \sigma_r^2 \boldsymbol{X}' + v^{(0)} \boldsymbol{I}$, $\boldsymbol{F}_2 = (-P_s \boldsymbol{X}'' + \bar{\gamma}_e \sigma_r^2 \boldsymbol{X}') \hat{\boldsymbol{g}}^\dagger$ and $c_2 = \hat{\boldsymbol{g}}^T (-P_s \boldsymbol{X}'' + \bar{\gamma}_e \sigma_r^2 \boldsymbol{X}') \hat{\boldsymbol{g}}^\dagger + \bar{\gamma}_e \sum_{k=1}^K \boldsymbol{g}_k^T \boldsymbol{Q}_k \boldsymbol{g}_k^\dagger + \bar{\gamma}_e^2 - v^{(0)} N_t \epsilon_0$. (4.46a) has been equivalently reformulated into (F.1). Then, given $\boldsymbol{g}_k$ fixed, applying similar procedure to that in Appendix E, it follows that there exists $v^{(1)} \geq 0$ such that the following LMI holds:

$$\begin{bmatrix} \boldsymbol{H}_2^{(1)} & \boldsymbol{F}_2{}^{(1)} & \boldsymbol{G}_2{}^{(1)} \\ \boldsymbol{F}_2{}^{(1)H} & c_2^{(1)} & \boldsymbol{B}_2{}^{(1)H} \\ \boldsymbol{G}_2{}^{(1)H} & \boldsymbol{B}_2{}^{(1)} & \boldsymbol{A}_2{}^{(1)} \end{bmatrix} - v^{(1)} \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \frac{-\boldsymbol{I}}{\epsilon_1} \end{bmatrix} \succeq \boldsymbol{0}. \tag{F.2}$$

Since $\frac{v^{(1)} \boldsymbol{I}}{\epsilon_1} + \boldsymbol{A}_2{}^{(1)} \succ \boldsymbol{0}$ always holds, (F.2) is equivalent to the following LMI:

$$\begin{bmatrix} \boldsymbol{A}_2{}^{(1)} + \frac{v^{(1)} \boldsymbol{I}}{\epsilon_1} & \boldsymbol{G}_2{}^{(1)H} & \boldsymbol{B}_2{}^{(1)} \\ \boldsymbol{G}_2{}^{(1)} & \boldsymbol{H}_2{}^{(1)} & \boldsymbol{F}_2{}^{(1)} \\ \boldsymbol{B}_2{}^{(1)H} & \boldsymbol{F}_2{}^{(1)H} & c_2^{(1)} - v^{(1)} \end{bmatrix} \succeq \boldsymbol{0}. \tag{F.3}$$

Next, devising the method of mathematical induction again as that for (E.3), (F.1) holds for $\boldsymbol{g}_k \in \mathcal{G}_k$, $\forall k$, if and only if there exists $\{v(k) \geq 0\}$, such that (4.65) is satisfied, which completes the proof.

Appendix G

Proof of Lemma 5.5.1

Since the optimal solution to (P1.1-SDP) is proved to be optimal to (P1.1), the optimum value for (P1.1-SDP) is $\tau f_1(\tau) = H_1(\tau)$. Hence, we identify the property of $H_1(\tau)$ by investigating problem (P1.1-SDP), the Lagrangian of which is given by

$$\begin{aligned}
&\mathcal{L}(\chi) = \\
&\mathrm{tr}\Big(\big(P_s\tilde{\boldsymbol{h}}_{sd}^{\dagger}\tilde{\boldsymbol{h}}_{sd}^{T} - \lambda\boldsymbol{D}_{\hat{sd}} + \sum_{k=1}^{K}\theta_k(1/\tau-1)\boldsymbol{D}_{\hat{se},k} - \sum_{k=1}^{K}\theta_k P_s\tilde{\boldsymbol{h}}_{se,k}^{\dagger}\tilde{\boldsymbol{h}}_{se,k}^{T} - \boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{U}\boldsymbol{W}_0^{\frac{1}{2}} \\
&+ \boldsymbol{Y}_1\big)\hat{\boldsymbol{X}}_1\Big) + \mathrm{tr}\Big(\big(-\lambda\boldsymbol{h}_{rd}^{\dagger}\boldsymbol{h}_{rd}^{T} + \sum_{k=1}^{K}\theta_k(1/\tau-1)\boldsymbol{h}_{re,k}^{\dagger}\boldsymbol{h}_{re,k}^{T} - \boldsymbol{U} + \boldsymbol{Y}_2\big)\hat{\boldsymbol{S}}\Big) \\
&+ \big(-\lambda\sigma_{n_d}^2 + \sum_{k=1}^{K}\theta_k(1/\tau-1)\sigma_{n_e,k}^2 + \mathrm{tr}(\boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{U}\boldsymbol{W}_0^{\frac{1}{2}}) + \zeta\big)\xi + \lambda\tau, \qquad \text{(G.1)}
\end{aligned}$$

where χ denotes a tuple consisting of all the primal and dual variables. Specifically, $\boldsymbol{Y}_1$, $\boldsymbol{Y}_2$ and λ are Lagrangian multipliers associated with $\hat{\boldsymbol{X}}_1$, $\hat{\boldsymbol{S}}$ and the first constraint of problem (P1.1-SDP), respectively; $\{\theta_k\}$ are the dual variables associated with the SINR constraint for the kth eavesdropper, respectively; $\boldsymbol{U} = diag([u_i]_{i=1}^N)$ with each diagonal entry u_i denoting the dual variable associated with the per-relay power constraint; ζ is the Lagrangian multiplier associated with $\xi \geq 0$. In addition, $\boldsymbol{W}_0 = diag([\eta\bar{\alpha}_i P_s|h_{sr_i}|^2]_{i=1}^N)$. The KKT conditions for (G.1) are partially listed as

follows.

$$\boldsymbol{Y}_1 = -P_s\tilde{\boldsymbol{h}}_{sd}^{\dagger}\tilde{\boldsymbol{h}}_{sd}^{T} + \lambda\boldsymbol{D}_{\hat{sd}} - \sum_{k=1}^{K}\theta_k(1/\tau-1)\boldsymbol{D}_{\hat{se},k} + \sum_{k=1}^{K}\theta_k P_s\tilde{\boldsymbol{h}}_{se,k}^{\dagger}\tilde{\boldsymbol{h}}_{se,k}^{T} + \boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{U}\boldsymbol{W}_0^{\frac{1}{2}} \tag{G.2a}$$

$$\boldsymbol{Y}_2 = \lambda\boldsymbol{h}_{rd}^{\dagger}\boldsymbol{h}_{rd}^{T} - \sum_{k=1}^{K}\theta_k(1/\tau-1)\boldsymbol{h}_{re,k}^{\dagger}\boldsymbol{h}_{re,k}^{T} + \boldsymbol{U} \tag{G.2b}$$

$$\zeta = \lambda\sigma_{n_d}^2 - \sum_{k=1}^{K}\theta_k(1/\tau-1)\sigma_{n_e,k}^2 - \mathrm{tr}(\boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{U}\boldsymbol{W}_0^{\frac{1}{2}}) \tag{G.2c}$$

The associated dual problem is accordingly given by

$$\begin{aligned}&\text{(P1.1-SDP-dual)}:\\ &\quad\min_{\lambda,\{\theta_k\},\{u_i\}}\ \lambda\tau\\ &\qquad\texttt{s.t.}\ \theta_k\geq 0,\ \forall k,\ u_i\geq 0,\ \forall i,\\ &\qquad\qquad\lambda\geq 0,\ \zeta\geq 0,\ \boldsymbol{Y}_1\succeq\boldsymbol{0},\ \boldsymbol{Y}_2\succeq\boldsymbol{0},\end{aligned}$$

where $\boldsymbol{Y}_1$, $\boldsymbol{Y}_2$, and ζ are given by (G.2a),(G.2b), and (G.2c), respectively. Since it is easily verified that (P1.1-SDP) satisfies the Slater condition, the strong duality holds [84]. This implies that the dual optimum value given by (P1.1-SDP-dual) is exactly $H_1(\tau)$, which turns out to be a point-wise minimum of a family of affine functions and thus concave over $\tau\in[\tau_{\min,1},1]$ [84, pp. 80].

Appendix H

Proof of Proposition 5.5.1

The KKT conditions for (G.1) also yields the following complementary slackness.

$$\boldsymbol{Y}_1 \hat{\boldsymbol{X}}_1^* = \boldsymbol{0} \tag{H.1a}$$

$$\boldsymbol{Y}_2 \hat{\boldsymbol{S}}^* = \boldsymbol{0} \tag{H.1b}$$

Pre- and post-multiply $\boldsymbol{W}_0^{\frac{1}{2}}$ with the left hand side (LHS) and right hand side (RHS) of equation (G.2b), respectively, and substitute $\boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{U}^*\boldsymbol{W}_0^{\frac{1}{2}}$ into (G.2a), $\boldsymbol{Y}_1^*$ can be rewritten as

$$\begin{aligned}\boldsymbol{Y}_1 = &- P_s\tilde{\boldsymbol{h}}_{sd}^{\dagger}\tilde{\boldsymbol{h}}_{sd}^{T} + \lambda\boldsymbol{D}_{\hat{sd}} - \sum_{k=1}^{K}\theta_k(1/\tau-1)\boldsymbol{D}_{\hat{se},k} - \lambda\boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{h}_{rd}^{\dagger}\boldsymbol{h}_{rd}^{T}\boldsymbol{W}_0^{\frac{1}{2}} + \sum_{k=1}^{K}\theta_k P_s\cdot \\ &\tilde{\boldsymbol{h}}_{se,k}^{\dagger}\tilde{\boldsymbol{h}}_{se,k}^{T} + \boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{Y}_2\boldsymbol{W}_0^{\frac{1}{2}} + \sum_{k=1}^{K}\theta_k(1/\tau-1)\boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{h}_{re,k}^{\dagger}\boldsymbol{h}_{re,k}^{T}\boldsymbol{W}_0^{\frac{1}{2}}.\end{aligned} \tag{H.2}$$

Next, introducing the notation of $[\cdot]_{\text{offd}}$ to represent a square matrix with its diagonal entries removed, it follows from (G.2b) that

$$[\boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{Y}_2\boldsymbol{W}_0^{\frac{1}{2}} - \lambda\boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{h}_{rd}^{\dagger}\boldsymbol{h}_{rd}^{T}\boldsymbol{W}_0^{\frac{1}{2}} - \sum_{k=1}^{K}\theta_k(1/\tau-1)\boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{h}_{re,k}^{\dagger}\boldsymbol{h}_{re,k}^{T}\boldsymbol{W}_0^{\frac{1}{2}}]_{\text{offd}} = \boldsymbol{0}. \tag{H.3}$$

Appendix H. Proof of Proposition 5.5.1

By subtracting (H.3) from (H.2), $\boldsymbol{Y}_1$ can be rewritten as follows.

$$\boldsymbol{Y}_1 = -P_s\tilde{\boldsymbol{h}}_{sd}^{\dagger}\tilde{\boldsymbol{h}}_{sd}^{T} + \lambda\boldsymbol{D}_{\hat{sd}} - \sum_{k=1}^{K}\theta_k(1/\tau-1)\boldsymbol{D}_{\hat{se},k} + [\boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{Y}_2\boldsymbol{W}_0^{\frac{1}{2}}]_{\mathrm{d}} + \sum_{k=1}^{K}\theta_k P_s \cdot \tilde{\boldsymbol{h}}_{se,k}^{\dagger}\tilde{\boldsymbol{h}}_{se,k}^{T} - [\lambda\boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{h}_{rd}^{\dagger}\boldsymbol{h}_{rd}^{T}\boldsymbol{W}_0^{\frac{1}{2}} - \sum_{k=1}^{K}\theta_k(1/\tau-1)\boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{h}_{re,k}^{\dagger}\boldsymbol{h}_{re,k}^{T}\boldsymbol{W}_0^{\frac{1}{2}}]_{\mathrm{d}}, \tag{H.4}$$

where $[\cdot]_{\mathrm{d}}$ denotes a square matrix with only the diagonal remained. Observing that

$$\begin{aligned}\boldsymbol{D}_{\hat{sd}}^{-1}\boldsymbol{D}_{\hat{se}} =& [\boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{h}_{rd}^{\dagger}\boldsymbol{h}_{rd}^{T}\boldsymbol{W}_0^{\frac{1}{2}}]_{\mathrm{d}}^{-1}[\boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{h}_{re,k}^{\dagger}\boldsymbol{h}_{re,k}^{T}\boldsymbol{W}_0^{\frac{1}{2}}]_{\mathrm{d}} \\ =& diag\left([|h_{r_ie,k}|^2 \,/\, |h_{r_id}|^2]_{i=1}^{N}\right),\end{aligned} \tag{H.5}$$

we denote (H.5) by $\boldsymbol{R}_{ed,k}$, $\forall k$. $\boldsymbol{Y}_1^*$ can thus be finally recast as

$$\boldsymbol{Y}_1 = -P_s\tilde{\boldsymbol{h}}_{sd}^{\dagger}\tilde{\boldsymbol{h}}_{sd}^{T} + \boldsymbol{\Xi} + \sum_{k=1}^{K}\theta_k P_s\tilde{\boldsymbol{h}}_{se,k}^{\dagger}\tilde{\boldsymbol{h}}_{se,k}^{T}, \tag{H.6}$$

where

$$\boldsymbol{\Xi} = [\boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{Y}_2\boldsymbol{W}_0^{\frac{1}{2}}]_{\mathrm{d}} - ([\boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{h}_{rd}^{\dagger}\boldsymbol{h}_{rd}^{T}\boldsymbol{W}_0^{\frac{1}{2}}]_{\mathrm{d}} - \boldsymbol{D}_{\hat{sd}})(\lambda\boldsymbol{I} - \sum_{k=1}^{K}\theta_k(1/\tau-1)\boldsymbol{R}_{ed,k}). \tag{H.7}$$

In the following we show that $\boldsymbol{\Xi} + \sum_{k=1}^{K}\theta_k P_s\tilde{\boldsymbol{h}}_{se,k}^{\dagger}\tilde{\boldsymbol{h}}_{se,k}^{T}$ is a positive definite matrix. Note that since $\boldsymbol{\Xi}$ is a diagonal matrix, its definiteness is only determined by the signs of its diagonal entries, for which we commence with the discussion in three difference cases.

1) **Case I**: $\exists i$ such that $\lambda - \sum_{k=1}^{K}\theta_k(1/\tau-1)[\boldsymbol{R}_{ed,k}]_{i,i} < 0$

Since $[[\boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{h}_{rd}^{\dagger}\boldsymbol{h}_{rd}^{T}\boldsymbol{W}_0^{\frac{1}{2}}]_{\mathrm{d}} - \boldsymbol{D}_{\hat{sd}}]_{i,i} = \eta\bar{\alpha}_i P_s|h_{sr_i}|^2|h_{r_id}|^2(1 - \frac{(1-\bar{\alpha}_i)\sigma_{n_a}^2+\sigma_{n_c}^2}{(1-\bar{\alpha}_i)(|h_{sr_i}|^2P_s+\sigma_{n_a}^2)+\sigma_{n_c}^2}) > 0$, it follows from (H.7) that $[\boldsymbol{\Xi}]_{i,i} > 0$ in this case.

2) **Case II**: $\exists i$ such that $\lambda - \sum_{k=1}^{K}\theta_k(1/\tau-1)[\boldsymbol{R}_{ed,k}]_{i,i} > 0$

We have $[[\boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{Y}_2\boldsymbol{W}_0^{\frac{1}{2}}]_{\mathrm{d}}]_{i,i} - [\boldsymbol{W}_0]_{i,i}|h_{r_id}|^2(\lambda^* - \sum_{k=1}^{K}\theta_k(1/\tau-1)[\boldsymbol{R}_{ed,k}]_{i,i}) \geq 0$ in accordance with (G.2b), which implies that $[\boldsymbol{\Xi}]_{i,i} = [[\boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{Y}_2\boldsymbol{W}_0^{\frac{1}{2}}]_{\mathrm{d}}]_{i,i} -$

Appendix H. Proof of Proposition 5.5.1

$[\boldsymbol{W}_0]_{i,i}|h_{r_id}|^2(\lambda-\sum_{k=1}^K\theta_k(1/\tau-1)[\boldsymbol{R}_{ed,k}]_{i,i})+[\boldsymbol{D}_{\hat{sd}}]_{i,i}(\lambda-\sum_{k=1}^K\theta_k(1/\tau-1)[\boldsymbol{R}_{ed,k}]_{i,i})>0$ (c.f. (H.7))

3) **Case III**: $\exists i$ such that $\lambda-\sum_{k=1}^K\theta_k(1/\tau-1)[\boldsymbol{R}_{ed,k}]_{i,i}=0$

In this case, it follows that $[\boldsymbol{\Xi}]_{i,i}=[\boldsymbol{W}_0^{\frac{1}{2}}\boldsymbol{Y}_2\boldsymbol{W}_0^{\frac{1}{2}}]_{i,i}\geq 0$. It is noteworthy that the number of i such that $\lambda-\sum_{k=1}^K\theta_k(1/\tau-1)[\boldsymbol{R}_{ed,k}]_{i,i}=0$ cannot exceed one. This can be proved by contradiction. (If $\exists i_1,i_2$, $i_1\neq i_2$, such that $\lambda-\sum_{k=1}^K\theta_k(1/\tau-1)[\boldsymbol{R}_{ed,k}]_{i_1,i_1}=0$ and $\lambda-\sum_{k=1}^K\theta_k(1/\tau-1)[\boldsymbol{R}_{ed,k}]_{i_2,i_2}=0$, it implies that $\sum_{k=1}^K\theta_k[\boldsymbol{R}_{ed,k}]_{i_1,i_1}=\sum_{k=1}^K\theta_k[\boldsymbol{R}_{ed,k}]_{i_2,i_2}$, which contradicts to the fact that for any two independent continuously distributed RVs, the chance that they are equal is zero.)

In a summary, $[\boldsymbol{\Xi}]_{i,i}\geq 0$, $\forall i$. If $[\boldsymbol{\Xi}]_{i,i}>0$, $\forall i$, it is obvious that $\boldsymbol{\Xi}+\sum_{k=1}^K\theta_kP_s\tilde{\boldsymbol{h}}_{se,k}^{\dagger}\tilde{\boldsymbol{h}}_{se,k}^T\succ\boldsymbol{0}$. Next, we show that it still holds true in the case that $\exists i'$, such that $[\boldsymbol{\Xi}]_{i',i'}=0$, $i'\in\mathcal{N}$, by definition of positive definite matrix. Define the null-space of $\boldsymbol{\Xi}$ by $\psi=\{\boldsymbol{\eta}|\boldsymbol{\eta}=\alpha\boldsymbol{e}_{i'},\alpha\in\mathbb{C}\}$ and multiply $\boldsymbol{\eta}^H$ and $\boldsymbol{\eta}$, $\forall\boldsymbol{\eta}\neq\boldsymbol{0}$, on LHS and RHS of $\boldsymbol{\Xi}+\sum_{k=1}^K\theta_kP_s\tilde{\boldsymbol{h}}_{se,k}^{\dagger}\tilde{\boldsymbol{h}}_{se,k}^T$, respectively. If $\boldsymbol{\eta}\notin\psi$, it is straightforward to obtain $\boldsymbol{\eta}^H(\boldsymbol{\Xi}+\sum_{k=1}^K\theta_kP_s\tilde{\boldsymbol{h}}_{se,k}^{\dagger}\tilde{\boldsymbol{h}}_{se,k}^T)\boldsymbol{\eta}>0$; otherwise, it follows that $\boldsymbol{\eta}^H(\boldsymbol{\Xi}+\sum_{k=1}^K\theta_kP_s\tilde{\boldsymbol{h}}_{se,k}^{\dagger}\tilde{\boldsymbol{h}}_{se,k}^T)\boldsymbol{\eta}=\sum_{k=1}^K\theta_kP_s\alpha^2|[\tilde{\boldsymbol{h}}_{se,k}]_{i'}|^2>0$, since $[\tilde{\boldsymbol{h}}_{se,k}]_{i'}\neq 0$ in probability. As a result, $\boldsymbol{Y}_1$ in (H.6) is shown to always take on a special structure, that is, a full-rank matrix minus a rank-one matrix. Note that this observation plays a key role in proving rank-one $\hat{\boldsymbol{X}}_1^*$, which is also identified in [51, Appendix C].

Finally, multiplying both sides of (H.6) by $\hat{\boldsymbol{X}}_1^*$, as per (H.1a), we obtain that $\hat{\boldsymbol{X}}_1^*=P_s(\Xi+\sum_{k=1}^K\theta_kP_s\tilde{\boldsymbol{h}}_{se,k}^{\dagger}\tilde{\boldsymbol{h}}_{se,k}^T)^{-1}\tilde{\boldsymbol{h}}_{sd}^{\dagger}\tilde{\boldsymbol{h}}_{sd}^T\hat{\boldsymbol{X}}_1^*$, which further implies that $\text{rank}(\hat{\boldsymbol{X}}_1^*)\leq\text{rank}(\tilde{\boldsymbol{h}}_{sd}^{\dagger}\tilde{\boldsymbol{h}}_{sd}^T)=1$. In addition, since the optimality of (P1.1-SDP) suggests that $\hat{\boldsymbol{X}}_1^*\neq\boldsymbol{0}$, $\text{rank}(\hat{\boldsymbol{X}}_1^*)=1$ is thus proved.

As $\hat{\boldsymbol{X}}_1^*$ can be decomposed as $\hat{\boldsymbol{w}}_1^*\hat{\boldsymbol{w}}_1^{*H}$ by EVD, (H.1a) results in $\boldsymbol{Y}_1\hat{\boldsymbol{w}}_1^*=\boldsymbol{0}$, which further implies that

$$\hat{\boldsymbol{w}}_1^*=P_s(\boldsymbol{\Xi}+\sum_{k=1}^K\theta_kP_s\tilde{\boldsymbol{h}}_{se,k}^{\dagger}\tilde{\boldsymbol{h}}_{se,k}^T)^{-1}\tilde{\boldsymbol{h}}_{sd}^{\dagger}\tilde{\boldsymbol{h}}_{sd}^T\hat{\boldsymbol{w}}_1^*. \tag{H.8}$$

(H.8) admits a unique solution $\hat{\boldsymbol{w}}_1$ up to a scaling factor, which is given by

$$\hat{\boldsymbol{w}}_1 = (\boldsymbol{\Xi} + \sum_{k=1}^{K} \theta_k P_s \tilde{\boldsymbol{h}}_{se,k}^{\dagger} \tilde{\boldsymbol{h}}_{se,k}^{T})^{-1} \tilde{\boldsymbol{h}}_{sd}^{\dagger}. \tag{H.9}$$

Consequently, we have $\hat{\boldsymbol{w}}_1^* = \beta \hat{\boldsymbol{w}}_1$, where $\beta \in \mathbb{R}_+$. On the other hand, by plugging $\hat{\boldsymbol{w}}_1^* = \beta \hat{\boldsymbol{w}}_1$ into the equality constraint of (P1.1-SDP), we have $\beta = \sqrt{\frac{\tau - \xi^* \sigma_{n_d}^2 - \mathrm{tr}(\hat{\boldsymbol{S}}^* \boldsymbol{h}_{rd}^{\dagger} \boldsymbol{h}_{rd}^{T})}{\mathrm{tr}(\hat{\boldsymbol{w}}_1 \hat{\boldsymbol{w}}_1^H \boldsymbol{D}_{\hat{s}d})}}$ and $\hat{\boldsymbol{w}}_1^* = \sqrt{\frac{\tau - \xi^* \sigma_{n_d}^2 - \mathrm{tr}(\hat{\boldsymbol{S}}^* \boldsymbol{h}_{rd}^{\dagger} \boldsymbol{h}_{rd}^{T})}{\mathrm{tr}(\hat{\boldsymbol{w}}_1 \hat{\boldsymbol{w}}_1^H \boldsymbol{D}_{\hat{s}d})}} \hat{\boldsymbol{w}}_1$.

At last, we show part 3) of Proposition 5.5.1. For the case of $K \geq N$, it is obvious that $\mathrm{rank}(\hat{\boldsymbol{S}})^* \leq N$. For the case of $K < N$, we first prove that $\lambda \boldsymbol{h}_{rd}^{\dagger} \boldsymbol{h}_{rd}^{T} + \boldsymbol{U}$ is a full-rank matrix, i.e., $\lambda \boldsymbol{h}_{rd}^{\dagger} \boldsymbol{h}_{rd}^{T} + \boldsymbol{U} \succ \boldsymbol{0}$, by definition. According to (G.2b), $\boldsymbol{Y}_2$ can be rewritten as follows.

$$\boldsymbol{Y}_2 = (1+\lambda) \boldsymbol{h}_{rd}^{\dagger} \boldsymbol{h}_{rd}^{T} + \boldsymbol{U} - (\sum_{k=1}^{K} \theta_k (1/\tau - 1) \boldsymbol{h}_{re,k}^{\dagger} \boldsymbol{h}_{re,k}^{T} + \boldsymbol{h}_{rd}^{\dagger} \boldsymbol{h}_{rd}^{T}) \tag{H.10}$$

To facilitate the proof in the sequel, we also define a set Φ as $\Phi = \bigcap_{k \in \Omega} \mathbf{null}(\boldsymbol{h}_{re,k}^{T})$, where $\Omega = \{k \in \mathcal{K} | \theta_k > 0\}$.

Next, multiply $\boldsymbol{\eta}^H$ and $\boldsymbol{\eta}$, $\boldsymbol{\eta} \neq \boldsymbol{0}$, with LHS and RHS of $\boldsymbol{Y}_2$ (c.f. (H.10)), respectively and discuss the following cases based on $\boldsymbol{\eta}$'s relation with $\mathbf{null}(\boldsymbol{Y}_2)$ and/or Φ.

1) **Case I**: $\boldsymbol{\eta} \in \mathbf{null}(\boldsymbol{Y}_2) \setminus \Phi$

It follows that

$$\boldsymbol{\eta}^H ((1+\lambda) \boldsymbol{h}_{rd}^{\dagger} \boldsymbol{h}_{rd}^{T} + \boldsymbol{U}) \boldsymbol{\eta} \overset{(a)}{=} \boldsymbol{\eta}^H (\sum_{k=1}^{K} \theta_k (1/\tau - 1) \boldsymbol{h}_{re,k}^{\dagger} \boldsymbol{h}_{re,k}^{T} + \boldsymbol{h}_{rd}^{\dagger} \boldsymbol{h}_{rd}^{T}) \boldsymbol{\eta} \overset{(b)}{>} 0, \tag{H.11}$$

where (a) is attained by $\boldsymbol{\eta}^H \boldsymbol{Y}_2 \boldsymbol{\eta} = 0$ (c.f. (H.10)) and (b) is due to the fact that either $\exists k$ such that $|\boldsymbol{\eta}^H \boldsymbol{h}_{re,k}^{\dagger}|^2 > 0$ or $|\boldsymbol{\eta}^H \boldsymbol{h}_{rd}^{\dagger}|^2 > 0$, otherwise $\boldsymbol{\eta} \in \Phi$.

2) **Case II**: $\boldsymbol{\eta} \notin \mathbf{null}(\boldsymbol{Y}_2)$

Together with $\boldsymbol{Y}_2 \succeq \boldsymbol{0}$, $\boldsymbol{\eta}^H \boldsymbol{Y}_2 \boldsymbol{\eta} > 0$ holds and substituting (H.10) for $\boldsymbol{Y}_2$, it follows

that

$$\boldsymbol{\eta}^H((1+\lambda)\boldsymbol{h}_{rd}^{\dagger}\boldsymbol{h}_{rd}^T+\boldsymbol{U})\boldsymbol{\eta} > \boldsymbol{\eta}^H(\sum_{k=1}^{K}\theta_k(1/\tau-1)\boldsymbol{h}_{re,k}^{\dagger}\boldsymbol{h}_{re,k}^T+\boldsymbol{h}_{rd}^{\dagger}\boldsymbol{h}_{rd}^T)\boldsymbol{\eta}\geq 0. \quad \text{(H.12)}$$

3) **Case III**: $\boldsymbol{\eta}\in\mathbf{null}(\boldsymbol{Y}_2)\cap\Phi\neq\emptyset$

As a result of (H.1b), the singular vectors of $\hat{\boldsymbol{S}}^*$ is composed of any orthogonal set of $\mathbf{null}(\boldsymbol{Y}_2)$, Hence, in this case $\exists\hat{\boldsymbol{S}}^*$ such that $\hat{\boldsymbol{S}}^*=\sum_{i=1}^{r}\alpha_i\boldsymbol{\eta}_i\boldsymbol{\eta}_i^H+\beta\boldsymbol{\eta}\boldsymbol{\eta}^H$, $\beta>0$, where $\{\boldsymbol{\eta}_i\}$, $i=1,\ldots,r$, $\boldsymbol{\eta}$ are orthogonal each other, and $r=\mathrm{rank}(\hat{\boldsymbol{S}}^*)$. As such, we reconstruct $\hat{\boldsymbol{S}}$ that is given by $\hat{\boldsymbol{S}}'=\hat{\boldsymbol{S}}^*-\beta\boldsymbol{\eta}\boldsymbol{\eta}^H$. Since $\boldsymbol{\eta}\in\Phi$, i.e., $\boldsymbol{h}_{re,k}^T\boldsymbol{\eta}=0$, $k\in\Omega$, and $\boldsymbol{h}_{rd}^T\boldsymbol{\eta}=0$, it is easy to verify that $\hat{\boldsymbol{S}}'$ satisfies all the constraints of problem (P1.1-SDP). Furthermore, the per-relay power constraints w.r.t. $\hat{\boldsymbol{X}}_1$ turn out to be

$$\mathrm{tr}((\hat{\boldsymbol{S}}^*+\eta\bar{\alpha}_iP_s|h_{sr_i}|^2\hat{\boldsymbol{X}}_1)\boldsymbol{E}_i)\leq\xi\eta\bar{\alpha}_iP_s|h_{sr_i}|^2+\beta\mathrm{tr}(\boldsymbol{\eta}\boldsymbol{\eta}^H\boldsymbol{E}_i),\ \forall i, \quad \text{(H.13)}$$

which enlarge the feasible region in terms of $\hat{\boldsymbol{X}}_1$ compared with that given the original $\hat{\boldsymbol{S}}^*$. It thus implies that given $\hat{\boldsymbol{S}}'$, problem (P1.1-SDP) yields a larger optimum value in general, which violate the optimality of $\hat{\boldsymbol{S}}'$. This contradiction is caused by the assumption of **Case III**, which shows that $\boldsymbol{\eta}\in\mathbf{null}(\boldsymbol{Y}_2)\cap\Phi=\emptyset$.

In a summary, since there exist only **Case I** and **Case II**, we have shown that $\forall\boldsymbol{\eta}\neq\mathbf{0}$, $\boldsymbol{\eta}^H((1+\lambda)\boldsymbol{h}_{rd}^{\dagger}\boldsymbol{h}_{rd}^T+\boldsymbol{U})\boldsymbol{\eta}>0$ holds, and therefore $\lambda\boldsymbol{h}_{rd}^{\dagger}\boldsymbol{h}_{rd}^T+\boldsymbol{U}$ is also a full-rank matrix. Finally, according to (G.2b), $\mathrm{rank}(\boldsymbol{Y}_2)\geq N-K$ since $\mathrm{rank}(\sum_{k=1}^{K}\theta_k(1/\tau-1)\boldsymbol{h}_{re,k}^{\dagger}\boldsymbol{h}_{re,k}^T)\leq K$. Combining with (H.1b), $\mathrm{rank}(\hat{\boldsymbol{S}}^*)\leq K$ is thus obtained.

Appendix I

Proof of Proposition 5.5.2

First, we exploit the following lemma to rewrite (P2′.1-SDR).

Lemma I.0.1. *Problem* (P2′.1-SDR) *is equivalent to the following problem.*

$$\text{(P2′.1-SDR-Eqv)}: \begin{cases} \underset{\boldsymbol{U}_1, \boldsymbol{U}_{2,\{x_i\},\{y_i\}}}{\texttt{max}} & (5.41) \\ \texttt{s.t.} & (5.42),\ \forall k,\ \text{(I.1)},\ \text{(I.2)},\ \forall i, \\ & \mathrm{tr}(\boldsymbol{U}_1\boldsymbol{E}_i) \le x_i,\ \mathrm{tr}(\boldsymbol{U}_2\boldsymbol{E}_i) \ge y_i,\ \forall i, \\ & \mathrm{tr}((\boldsymbol{U}_1 - \boldsymbol{U}_2)\boldsymbol{E}_i) \le 0,\ \forall i, \\ & \boldsymbol{U}_1 \succeq \boldsymbol{0},\ \boldsymbol{U}_2 \succeq \boldsymbol{0}, \end{cases}$$

where (I.1) *and* (I.2) *are given by*

$$\left\| \begin{array}{c} 2\sigma_{n_c}\mathrm{tr}(\boldsymbol{U}_2\boldsymbol{E}_i) \\ 2\sqrt{\left(1 - \frac{\bar{z}_i}{c_{0,i}}\right)\frac{1}{c_{1,i}}} \\ \left(1 - \frac{\bar{z}_i}{c_{0,i}} - c_{1,i}\mathrm{tr}(\boldsymbol{U}_1\boldsymbol{E}_i)\right) - \left(\mathrm{tr}(\boldsymbol{U}_2\boldsymbol{E}_i) + \frac{1}{c_{1,i}}\right) \end{array} \right\| \le \left(1 - \frac{\bar{z}_i}{c_{0,i}} - c_{1,i}x_i\right) + \left(y_i + \frac{1}{c_{1,i}}\right) \quad \text{(I.1)}$$

and

$$\left\| \begin{array}{c} 2\sigma_{n_c}\mathrm{tr}(\boldsymbol{U}_2\boldsymbol{E}_i) \\ 2\sqrt{\frac{1}{c_{1,i}}} \\ (1 - c_{1,i}\mathrm{tr}(\boldsymbol{U}_1\boldsymbol{E}_i)) - \left(\mathrm{tr}(\boldsymbol{U}_2\boldsymbol{E}_i) + \frac{1}{c_{1,i}}\right) \end{array} \right\| \le (1 - c_{1,i}x_i) + \left(y_i + \frac{1}{c_{1,i}}\right), \quad \text{(I.2)}$$

respectively.

Proof. For the convenience of the proof, the optimum value for (P2′.1-SDR)

and (P2′.1-SDR-Eqv) are denoted by f_0^* and $\tilde{f}_0^*$, respectively. Assuming that $(\boldsymbol{U}_1^*, \boldsymbol{U}_2^*, \{x_i^*\}, \{y_i^*\})$ is the optimal solution to (P2′.1-SDR), it is easily verified to be feasible for (P2′.1-SDR-Eqv) as well, which implies that $f_0^* \leq \tilde{f}_0^*$. On the other hand, if problem (P2′.1-SDR-Eqv) returns an optimal solution of $(\tilde{\boldsymbol{U}}_1^*, \tilde{\boldsymbol{U}}_2^*, \{\tilde{x}_i^*\}, \{\tilde{y}_i^*\})$, by defining $\mathrm{tr}(\tilde{\boldsymbol{U}}_1^* \boldsymbol{E}_i) = x_i'^*$ and $\mathrm{tr}(\tilde{\boldsymbol{U}}_2^* \boldsymbol{E}_i) = y_i'^*$, $\forall i$, we can show that $(\tilde{\boldsymbol{U}}_1^*, \tilde{\boldsymbol{U}}_2^*, \{x_i'^*\}, \{y_i'^*\})$ is also feasible for (P2′.1-SDR) as follows. As for (5.37),

$$\begin{aligned}
\left\| \begin{array}{c} 2\sigma_{n_c} y_i'^* \\ 2\sqrt{\left(1 - \frac{\bar{z}_i}{c_{0,i}}\right)\frac{1}{c_{1,i}}} \\ \left(1 - \frac{\bar{z}_i}{c_{0,i}} - c_{1,i} x_i'^*\right) - \left(y_i'^* + \frac{1}{c_{1,i}}\right) \end{array} \right\| &\overset{(a)}{\leq} \left(1 - \frac{\bar{z}_i}{c_{0,i}} - c_{1,i}\tilde{x}_i^*\right) + \left(\tilde{y}_i^* + \frac{1}{c_{1,i}}\right) \\
&\overset{(b)}{\leq} \left(1 - \frac{\bar{z}_i}{c_{0,i}} - c_{1,i}\mathrm{tr}(\tilde{\boldsymbol{U}}_1^* \boldsymbol{E}_i)\right) + \left(\mathrm{tr}(\tilde{\boldsymbol{U}}_2^* \boldsymbol{E}_i) + \frac{1}{c_{1,i}}\right) \\
&= \left(1 - \frac{\bar{z}_i}{c_{0,i}} - c_{1,i} x_i'^*\right) + \left(y_i'^* + \frac{1}{c_{1,i}}\right), \qquad \text{(I.3)}
\end{aligned}$$

in which (a) is due to (I.1), and (b) comes from $\mathrm{tr}(\tilde{\boldsymbol{U}}_1^* \boldsymbol{E}_i) \leq \tilde{x}_i^*$ and $\mathrm{tr}(\tilde{\boldsymbol{U}}_2^* \boldsymbol{E}_i) \geq \tilde{y}_i^*$. Similarly, $(\tilde{\boldsymbol{U}}_1^*, \tilde{\boldsymbol{U}}_2^*, \{x_i'^*\}, \{y_i'^*\})$ can also be proved to satisfy (5.39). In addition, $x_i'^* - y_i'^* = \mathrm{tr}(\tilde{\boldsymbol{U}}_1^* \boldsymbol{E}_i) - \mathrm{tr}(\tilde{\boldsymbol{U}}_2^* \boldsymbol{E}_i) \leq 0$, $\forall i$, i.e., (5.40) holds true. These feasibility implies that $\tilde{f}_0^* \leq f_0^*$. By combining the above two facts, we have $\tilde{f}_0^* = f_0^*$, which completes the proof. □

Then, we apply the Charnes-Cooper transformation again to (P2′.1-SDR-Eqv), the result of which is denoted by (P2′.1-SDP-Eqv). It is noteworthy that the Charnes-Cooper transformed constraint of (I.1) admits the form given by $\|\boldsymbol{x}^{(i)}\| \leq h(\hat{x}_i, \hat{y}_i)$, $\forall i$, where $\boldsymbol{x}^{(i)}$ is the column vector inside $\|\cdot\|$ of the LHS of (I.1) and $h(\hat{x}_i, \hat{y}_i)$ indicates the RHS. Since it is easy to check that $\xi > 0$ as a result of feasibility, we have $\|\boldsymbol{x}^{(i)}\| > 0 \Rightarrow h(\hat{x}_i, \hat{y}_i) > 0$, which implies that

$$\begin{bmatrix} h(\hat{x}_i, \hat{y}_i) & \boldsymbol{x}^{(i)H} \\ \boldsymbol{x}^{(i)} & h(\hat{x}_i, \hat{y}_i)\boldsymbol{I} \end{bmatrix} \succeq \boldsymbol{0} \qquad \text{(I.4)}$$

according to Schur Complement. (I.4) thus holds true, $\forall \boldsymbol{x}^{(i)}$ such that $\|\boldsymbol{x}^{(i)}\| \leq$

$h(\hat{x}_i, \hat{y}_i)$. Consequently, it enables us to show that (I.1) can be recast into a constraint not related to $\hat{\boldsymbol{U}}_1$, $\hat{\boldsymbol{U}}_2$, following the same procedure as [41, Appendix III] by exploiting [113, *Lemma 2*]. Similarly, the Charnes-Cooper transformed constraint of (I.2) can also be rewritten without $\hat{\boldsymbol{U}}_1$, $\hat{\boldsymbol{U}}_2$. Hence, the partial Lagrangian for (P2′.1-SDP-Eqv) in terms of $\hat{\boldsymbol{U}}_1$ and $\hat{\boldsymbol{U}}_2$ can be expressed as

$$\begin{aligned}\mathcal{L}(\varphi) =&\mathrm{tr}\Big(\big(P_s\boldsymbol{s}_{sd}^{\dagger}\boldsymbol{s}_{sd}^{T} - \lambda\sigma_{n_a}^2 diag(\boldsymbol{c}_0 \circ \|\boldsymbol{h}_{rd}\|.^2) + \sum_{k=1}^{K}\theta_k(1/\tau-1)\sigma_{n_a}^2 diag(\boldsymbol{c}_0 \circ \|\boldsymbol{h}_{re,k}\|.^2)\\ &- P_s\sum_{k=1}^{K}\theta_k\boldsymbol{s}_{se,k}^{\dagger}\boldsymbol{s}_{se,k}^{T} - \boldsymbol{\Delta} - \boldsymbol{\Sigma} + \boldsymbol{Y}_1\big)\hat{\boldsymbol{U}}_1\Big) + \mathrm{tr}\Big(\big(-\lambda\sigma_{n_c}^2 diag(\boldsymbol{c}_0 \circ \|\boldsymbol{h}_{rd}\|.^2)\\ &+ \sum_{k=1}^{K}\theta_k(1/\tau-1)\sigma_{n_c}^2 diag(\boldsymbol{c}_0 \circ \|\boldsymbol{h}_{re,k}\|.^2) + \boldsymbol{\Pi} + \boldsymbol{\Sigma} + \boldsymbol{Y}_2\big)\hat{\boldsymbol{U}}_2\Big) + \lambda\tau, \end{aligned} \tag{I.5}$$

where φ denotes a tuple comprising all the associated primal and dual variables: $\boldsymbol{Y}_1$, $\boldsymbol{Y}_2$, and $\{\theta_k\}$ are Lagrangian multipliers associated with $\hat{\boldsymbol{U}}_1$, $\hat{\boldsymbol{U}}_2$, and (5.42), $\forall k$, respectively; λ is the dual variable associated with the only equality constraint; $\boldsymbol{\Delta} = diag([\delta_i]_{i=1}^N)$ and $\boldsymbol{\Pi} = diag([\pi_i]_{i=1}^N)$ denote those associated with $\mathrm{tr}(\boldsymbol{U}_1\boldsymbol{E}_i) \le x_i$ and $\mathrm{tr}(\boldsymbol{U}_2\boldsymbol{E}_i) \ge y_i$, $\forall i$, respectively; finally, the diagonal entry of $\boldsymbol{\Sigma} = diag([\sigma_i]_{i=1}^N)$ denotes the dual variable associated with $\mathrm{tr}((\boldsymbol{U}_1 - \boldsymbol{U}_2)\boldsymbol{E}_i) \le 0$, $\forall i$. The KKT conditions related to (I.5) are accordingly given by

$$\boldsymbol{Y}_1 = -P_s\boldsymbol{s}_{sd}^{\dagger}\boldsymbol{s}_{sd}^{T} + \boldsymbol{\Xi}' + P_s\sum_{k=1}^{K}\theta_k\boldsymbol{s}_{se,k}^{\dagger}\boldsymbol{s}_{se,k}^{T}, \tag{I.6a}$$

$$\boldsymbol{Y}_2 = \boldsymbol{D} - \boldsymbol{\Pi} - \boldsymbol{\Sigma}, \tag{I.6b}$$

$$\boldsymbol{Y}_1\hat{\boldsymbol{U}}_1^{*} = \boldsymbol{0}, \tag{I.6c}$$

$$\boldsymbol{Y}_2\hat{\boldsymbol{U}}_2^{*} = \boldsymbol{0}, \tag{I.6d}$$

where we introduce $\boldsymbol{\Xi}' = \frac{\sigma_{n_a}^2}{\sigma_{n_c}^2}\boldsymbol{D} + \boldsymbol{\Sigma} + \boldsymbol{\Delta}$, and $\boldsymbol{D} = \lambda\sigma_{n_c}^2 diag(\boldsymbol{c}_0 \circ \|\boldsymbol{h}_{rd}\|.^2) - \sum_{k=1}^{K}\theta_k(1/\tau-1)\sigma_{n_c}^2 diag(\boldsymbol{c}_0 \circ \|\boldsymbol{h}_{re,k}\|.^2)$ for the notation simplicity.

Next, we show that $\boldsymbol{\Xi}' + P_s\sum_{k=1}^{K}\theta_k\boldsymbol{s}_{se,k}^{\dagger}\boldsymbol{s}_{se,k}^{T}$ in (I.6a) is a positive definite matrix in the following two cases.

Appendix I. Proof of Proposition 5.5.2

1) **Case I**: $\theta_k = 0$, $\forall k \in \mathcal{K}$

In this case, since $\lambda > 0$ (c.f. (I.5)) due to the strong duality, it is easily verified that $\boldsymbol{D} \succ \mathbf{0}$ and therefore $\boldsymbol{\Xi}' \succ \mathbf{0}$.

2) **Case II**: $\exists k$ such that $\theta_k \neq 0$

Since $\boldsymbol{\Xi}'$ is a positive semidefinite diagonal matrix, it is shown to have maximum one zero diagonal entry according to the similar argument made in **Case III** of Appendix H (c.f. pp 165), and the positive definiteness of $\boldsymbol{\Xi}' + P_s \sum_{k=1}^{K} \theta_k \boldsymbol{s}_{se,k}^{\dagger} \boldsymbol{s}_{se,k}^{T} \succ \mathbf{0}$ can thus be proved by definition without difficulty. As $\boldsymbol{Y}_1$ (c.f. (I.6a)) again complies with the difference between a positive definite matrix and a rank one matrix, it turns out that $\text{rank}(\hat{\boldsymbol{U}}_1) \leq 1$ according to (I.6c). Then, following the same procedure as that in Appendix H, 2) of Proposition 5.5.2 can be proved with $\hat{\boldsymbol{u}}_1 = (\boldsymbol{\Xi}' + \sum_{k=1}^{K} \theta_k^* P_s \boldsymbol{s}_{se,k}^{\dagger} \boldsymbol{s}_{se,k}^{T})^{-1} \boldsymbol{s}_{sd}^{\dagger}$, the detailed of which is omitted here for brevity.

Finally, it is verified that (P2′.1-SDP) is related to $\hat{\boldsymbol{U}}_2^*$ merely with its diagonal entries, viz, $[\hat{\boldsymbol{U}}_2^*]_{i,i}$, $\forall i \in \mathcal{N}$, (c.f. (5.41), (5.42)). Furthermore, denoting $[[\hat{\boldsymbol{U}}_2^*]_{i,i}^{1/2}]_{i=1}^{N}$ by $\hat{\boldsymbol{u}}_2^*$, it is easily checked that the diagonal entries remain the same after we replace $\hat{\boldsymbol{U}}_2^*$ by $\hat{\boldsymbol{u}}_2^* \hat{\boldsymbol{u}}_2^{*H}$. Hence, we arrive at the conclusion that such modification returns a rank-one $\hat{\boldsymbol{U}}_2^*$ for (P2′.1-SDP), which completes the proof for 3).

Appendix J

Proof of Lemma 5.8.1

To facilitate the asymptotic analysis in the sequel, we first reexpress $\mathrm{SINR}_{\mathrm{S,D}}$ (c.f. (5.11)) and $\mathrm{SINR}_{\mathrm{S,E,k}}$ (c.f. (5.12)) using the rewritten channels as follows.

$$\mathrm{SINR}_{\mathrm{S,D}} = \frac{\underbrace{P_s|\sum_{i=1}^{N}\sqrt{\beta_{r_id}}\sqrt{\beta_{sr_i}}|\bar{h}_{r_id}|\bar{h}_{sr_i}|\sqrt{1-\alpha_i}|\beta_i|\,|^2}_{\mathrm{I}}}{\underbrace{\sigma_{n_a}^2\sum_{i=1}^{N}\beta_{r_id}|\bar{h}_{r_id}|^2|\beta_i|^2(1-\alpha_i)}_{\mathrm{II}}+\underbrace{\sigma_{n_c}^2\sum_{i=1}^{N}\beta_{r_id}|\bar{h}_{r_id}|^2|\beta_i|^2}_{\mathrm{III}}+\sigma_{n_d}^2} \tag{J.1}$$

$$\mathrm{SINR}_{\mathrm{S,E,k}} = \frac{\underbrace{P_s\,|\boldsymbol{h}_{re,k}^T\boldsymbol{D}_{\beta\alpha}\boldsymbol{h}_{sr}|^2}_{\mathrm{IV}}}{\underbrace{\sigma_{n_a}^2\sum_{i=1}^{N}\beta_{r_ie,k}|\bar{h}_{r_ie,k}|^2|\beta_i|^2(1-\alpha_i)}_{\mathrm{V}}+\underbrace{\sigma_{n_c}^2\sum_{i=1}^{N}\beta_{r_ie,k}|\bar{h}_{r_ie,k}|^2|\beta_i|^2}_{\mathrm{VI}}+\sigma_{n_e,k}^2} \tag{J.2}$$

Next, we examine each terms in (J.1) and (J.2).

To proceed with term I, by substituting β_i with (5.3), where $\measuredangle\beta_i = -\measuredangle\bar{h}_{r_id} - \measuredangle\bar{h}_{sr_i}$, we calculate the variance of each indexed part of the summation, i.e., X_n, $n = 1, \ldots, N$, which is a RV w.r.t. $|h_{sr_n}|$ and $|h_{r_nd}|$, given by

$$X_n = \sqrt{\beta_{r_nd}}\sqrt{\beta_{sr_n}}|\bar{h}_{r_nd}|\sqrt{\frac{\eta\alpha_n(1-\alpha_n)P_s\beta_{sr_n}|\bar{h}_{sr_n}|^4}{(1-\alpha_n)P_s\beta_{sr_n}|\bar{h}_{sr_n}|^2+(1-\alpha_n)\sigma_{n_a}^2+\sigma_{n_c}^2}}. \tag{J.3}$$

Appendix J. Proof of Lemma 5.8.1

To obtain $\mathsf{Var}[X_n] = \mathbb{E}[X_n^2] - \mathbb{E}^2[X_n]$, we derive $\mathbb{E}[X_n^2]$ as follows.

$$\begin{aligned}\mathbb{E}[X_n^2] &\overset{(a)}{=} \beta_{r_n d}\beta_{sr_n}\mathbb{E}\left[\frac{\eta\alpha_n(1-\alpha_n)P_s\beta_{sr_n}|\bar{h}_{sr_n}|^4}{(1-\alpha_n)P_s\beta_{sr_n}|\bar{h}_{sr_n}|^2+(1-\alpha_n)\sigma_{n_a}^2+\sigma_{n_c}^2}\right]\\ &\leq \beta_{r_n d}\beta_{sr_n}\mathbb{E}[\eta\alpha_n|\bar{h}_{sr_n}|^2]\\ &= \beta_{r_n d}\beta_{sr_n}\eta\alpha_n \leq \beta_{r_n d}\beta_{sr_n}\eta, \end{aligned} \tag{J.4}$$

where the cause of (a) is that $|\bar{h}_{r_n d}|^2$'s follows exponential distribution with unit mean. Denoting $\mathsf{Var}(X_n)$ by σ_n^2, $n = 1, \ldots, N$, it is easy to verify that $\Sigma_{n=1}^{\infty}\frac{\sigma_n^2}{n^2} \leq \Sigma_{n=1}^{\infty}\frac{\mathbb{E}[X_n^2]}{n^2} \leq \eta \max_{n\in\mathcal{N}} \beta_{r_n d}\beta_{sr_n}\Sigma_{n=1}^{\infty}\frac{1}{n^2} = \frac{\pi}{6}\eta \max_{n\in\mathcal{N}} \beta_{r_n d}\beta_{sr_n}$. As a result, the condition for Kolmogorov's strong law of large number (SLLN) [114, **Theorem 8.1**] holds, which implies that

$$\begin{aligned}&\frac{1}{N}\mathrm{I} \xrightarrow[N\to\infty]{\text{a.s.}}\\ &\frac{1}{N}\sum_{i=1}^{N}\sqrt{\beta_{r_i d}}\sqrt{\beta_{sr_i}}\sqrt{\eta\alpha_i(1-\alpha_i)P_s\beta_{sr_i}}\mathbb{E}[|\bar{h}_{r_i d}|]\cdot\\ &\mathbb{E}\left[\frac{Y_i}{\sqrt{(1-\alpha_i)P_s\beta_{sr_i}Y_i+(1-\alpha_i)\sigma_{n_a}^2+\sigma_{n_c}^2}}\right], \end{aligned} \tag{J.5}$$

where Y_i is the RV representing $|\bar{h}_{sr_i}|^2$ that is drawn from exponential distribution with unit mean and takes on a value y_i. Since $|\bar{h}_{r_i d}|$'s are *i.i.d.* Rayleigh RVs independent from Y_i's with $\sqrt{\frac{\pi}{4}}$ mean, the remaining task is to derive the expectation of $\frac{Y_i}{\sqrt{(1-\alpha_i)P_s\beta_{sr_i}Y_i+(1-\alpha_i)\sigma_{n_a}^2+\sigma_{n_c}^2}}$, $i = 1, \ldots, N$, as shown below.

$$\begin{aligned}\mathbb{E}\left[\frac{Y_i}{\sqrt{(1-\alpha_i)P_s\beta_{sr_i}Y_i+(1-\alpha_i)\sigma_{n_a}^2+\sigma_{n_c}^2}}\right] &= \int_0^{\infty}\frac{y_i}{\sqrt{(1-\alpha_i)P_s\beta_{sr_i}(y_i+v_i)}}e^{-y_i}\mathrm{d}y_i\\ &= \frac{e^{v_i}}{\sqrt{(1-\alpha_i)P_s\beta_{sr_i}}}\int_{v_i}^{\infty}\frac{\tilde{y}_i-v_i}{\sqrt{\tilde{y}_i}}e^{-\tilde{y}_i}\mathrm{d}\tilde{y}_i\\ &= \frac{1}{\sqrt{(1-\alpha_i)P_s\beta_{sr_i}}}v_i^{\frac{1}{4}}e^{\frac{v_i}{2}}W_{-\frac{3}{4},-\frac{3}{4}}(v_i), \end{aligned} \tag{J.6}$$

where $v_i = \frac{(1-\alpha_i)\sigma_{n_a}^2+\sigma_{n_c}^2}{(1-\alpha_i)P_s\beta_{sr_i}}$, $y_i + v_i$ is denoted by $\tilde{y}_i$, and the last equation is in accordance with [111, 3.383(4)], where $W_{\lambda,\mu}(\cdot)$ denotes the Whittaker function

[111, 9.22], with parameters λ, μ ; u, ν, μ and β are replaced with v_i, $\frac{1}{2}$, 2 and 1, respectively. Consequently, I $\xrightarrow[N\to\infty]{\text{a.s.}} \sum_{i=1}^{N} e_i W_{-\frac{3}{4},-\frac{3}{4}}(v_i)$ is obtained, where $e_i = \frac{\sqrt{\pi}}{2}\sqrt{\beta_{sr_i}}\sqrt{\beta_{r_id}}\sqrt{\eta\alpha_i}v_i^{\frac{1}{4}}e^{\frac{v_i}{2}}$.

To deal with II, by substituting (5.3) for β_i, $\forall i$, it follows that

$$\text{II} = \sigma_{n_a}^2 \sum_{i=1}^{N} \beta_{r_id}|\bar{h}_{r_id}|^2 \eta\alpha_i \tfrac{Y_i}{Y_i+v_i}. \tag{J.7}$$

As that for I, we verify the condition for Kolmogorov's SLLN on (J.7) and therefore, considering $|\bar{h}_{r_id}|^2$'s are *i.i.d.* exponential RVs independent from Y_i's, obtain that

$$\text{II} \xrightarrow[N\to\infty]{\text{a.s.}} \sigma_{n_a}^2 \sum_{i=1}^{N} \beta_{r_id}\eta\alpha_i \mathbb{E}\left[\tfrac{Y_i}{Y_i+v_i}\right]. \tag{J.8}$$

Moreover, as $\mathbb{E}[\frac{Y_i}{Y_i+v_i}] = 1-\mathbb{E}[\frac{v_i}{Y_i+v_i}] = 1-v_i\int_0^\infty \frac{e^{-y_i}}{y_i+v_i}\mathrm{d}y_i \overset{(a)}{=} 1+v_ie^{v_i}\mathrm{E_i}(-v_i)$, where (a) is due to [111, 3.382(4)] ($\beta = v_i$, $\nu = -1$, and $\mu = 1$) and $\mathrm{E_i}(-v_i)$ is the exponential integral defined by $\int_{-\infty}^{-v_i}\frac{e^t}{t}\mathrm{d}t$, for $v_i > 0$, we arrive at II $\xrightarrow[N\to\infty]{\text{a.s.}} \sigma_{n_a}^2\sum_{i=1}^{N}\beta_{r_id}\eta\alpha_i(1 + v_ie^{v_i}\mathrm{E_i}(-v_i))$.

Then, applying the same procedure as that for II, it is easy to obtain that

$$\text{III} \xrightarrow[N\to\infty]{\text{a.s.}} \sigma_{n_c}^2 \sum_{i=1}^{N} \beta_{r_id}\frac{\eta\alpha_i}{1-\alpha_i}\mathbb{E}\left[\tfrac{Y_i}{Y_i+v_i}\right], \tag{J.9}$$

and therefore III $\xrightarrow[N\to\infty]{\text{a.s.}}$ $\sigma_{n_c}^2\sum_{i=1}^{N}\beta_{r_id}\frac{\eta\alpha_i}{1-\alpha_i}(1 + v_ie^{v_i}\mathrm{E_i}(-v_i))$, since $\mathbb{E}[\frac{Y_i}{Y_i+v_i}] = 1 + v_ie^{v_i}\mathrm{E_i}(-v_i)$.

Combining I, II, and III, $\mathrm{SINR_{S,D}}$ is given by (c.f. (J.1))

$$\mathrm{SINR_{S,D}} = \frac{P_s|\sum_{i=1}^{N} e_i W_{-\frac{3}{4},-\frac{3}{4}}(v_i)|^2}{\sum_{i=1}^{N}\beta_{r_id}\eta\alpha_i(\sigma_{n_a}^2 + \frac{\sigma_{n_c}^2}{1-\alpha_i})(1 + v_ie^{v_i}\mathrm{E_i}(-v_i)) + \sigma_{n_d}^2}. \tag{J.10}$$

Identifying $|\boldsymbol{h}_{re,k}^T\boldsymbol{D}_{\beta\alpha}\boldsymbol{h}_{sr}|^2 = \left|\frac{\bar{\boldsymbol{h}}_{re,k}^H\boldsymbol{D}_{\beta_{re,k}}\boldsymbol{D}_{\beta\alpha}^\dagger\boldsymbol{h}_{sr}}{\|\boldsymbol{D}_{\beta_{re,k}}\boldsymbol{D}_{\beta\alpha}^\dagger\boldsymbol{h}_{sr}\|}\right|^2 \|\boldsymbol{D}_{\beta_{re,k}}\boldsymbol{D}_{\beta\alpha}^\dagger\boldsymbol{h}_{sr}\|^2$, IV can be recast as a Chi-square distributed RV with 2 d.o.f [107, *Lemma 1*], denoted by Z,

multiplied by $\|\boldsymbol{D}_{\beta_{re,k}}\boldsymbol{D}_{\beta\alpha}^{\dagger}\boldsymbol{h}_{sr}\|^2$, where $\boldsymbol{D}_{\beta_{re,k}} = diag([\sqrt{\beta_{r_ie,k}}]_{i=1}^{N})$, $\forall k \in \mathcal{K}$. Hence, the asymptotic expression for IV merely depends on $\|\boldsymbol{D}_{\beta_{re,k}}\boldsymbol{D}_{\beta\alpha}^{\dagger}\boldsymbol{h}_{sr}\|^2$ that is given by

$$\begin{aligned}\|\boldsymbol{D}_{\beta_{re,k}}\boldsymbol{D}_{\beta\alpha}^{\dagger}\boldsymbol{h}_{sr}\|^2 &= \sum_{i=1}^{N}\beta_{r_ie,k}\beta_{sr_i}\eta\alpha_i\tfrac{Y_i^2}{Y_i+v_i}\\ &\overset{(a)}{=} \sum_{i=1}^{N}\beta_{r_ie,k}\beta_{sr_i}\eta\alpha_i\mathbb{E}\left[\tfrac{Y_i^2}{Y_i+v_i}\right]\\ &\overset{(b)}{=} \sum_{i=1}^{N}\beta_{r_ie,k}\beta_{sr_i}\eta\alpha_i(1-v_i-v_i^2e^{v_i}\mathrm{E_i}(-v_i)),\end{aligned} \tag{J.11}$$

where (a) is as a result of SLLN and (b) is due to [111, 3.382(4)], the same as that for II. In accordance with (J.11), we obtain that IV $\xrightarrow[N\to\infty]{\text{a.s.}} Z\sum_{i=1}^{N}\beta_{r_ie,k}\beta_{sr_i}\eta\alpha_i(1-v_i-v_i^2e^{v_i}\mathrm{E_i}(-v_i))$. Similar to II and III, since $|\bar{h}_{r_ie,k}|^2$'s, $\forall k \in \mathcal{K}$, are also RVs independent from Y_i's, V and VI can be asymptotically expressed as $\sigma_{n_a}^2\sum_{i=1}^{N}\beta_{r_ie,k}\eta\alpha_i(1+v_ie^{v_i}\mathrm{E_i}(-v_i))$ and $\sigma_{n_c}^2\sum_{i=1}^{N}\beta_{r_ie,k}\frac{\eta\alpha_i}{1-\alpha_i}(1+v_ie^{v_i}\mathrm{E_i}(-v_i))$, respectively. Consequently, $\mathrm{SINR_{S,E,k}}$, substituting the above asymptotic expressions for IV, V, and VI, is given by (c.f. (J.2))

$$\mathrm{SINR_{S,E,k}} = \frac{P_sZ\sum_{i=1}^{N}\beta_{r_ie,k}\beta_{sr_i}\eta\alpha_i(1-v_i-v_i^2e^{v_i}\mathrm{E_i}(-v_i))}{\sum_{i=1}^{N}\beta_{r_ie,k}\eta\alpha_i(\sigma_{n_a}^2+\frac{\sigma_{n_c}^2}{1-\alpha_i})(1+v_ie^{v_i}\mathrm{E_i}(-v_i))+\sigma_{n_e,k}^2}. \tag{J.12}$$

Further, replacing the corresponding asymptotic expressions for $r_{\mathrm{S,D}}$ (c.f. (J.10)) and $r_{\mathrm{S,E,k}}$ (c.f. (J.12)) for $K=1$ in line with (5.13), Lemma 5.8.1 is proved.

References

[1] X. Zhou, L. Song, and Y. Zhang, *Physical Layer Security in Wireless Communications*. CRC Press, 2013, Chap. 1 Fundamentals of physical layer security.

[2] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, July 2006.

[3] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless networks with RF energy harvesting: A contemporary survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 757–789, Second Quart. 2015.

[4] S. Bi, C. K. Ho, and R. Zhang, "Wireless powered communication: opportunities and challenges," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 117–125, April 2015.

[5] P. Grover and A. Sahai, "Shannon meets Tesla: Wireless information and power transfer," in *Proc. IEEE International Symposium on Information (ISIT'10)*, Austin, TX, USA, June 2010, pp. 2363–2367.

[6] M. Bloch, J. Barros, M. R. D. Rodrigues, and S. McLaughlin, "Wireless information-theoretic security," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2515–2534, June 2008.

[7] E. Tekin and A. Yener, "The general Gaussian multiple-access and two-way wiretap channels: Achievable rates and cooperative jamming," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2735–2751, June 2008.

[8] L. Dong, Z. Han, A. P. Petropulu, and H. V. Poor, "Improving wireless physical layer security via cooperating relays," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1875–1888, March 2010.

[9] X. Zhou, R. Zhang, and C. Ho, "Wireless information and power transfer: architecture design and rate-energy tradeoff," *IEEE Trans.commun.*, vol. 61, no. 11, pp. 4757–4767, Nov. 2013.

[10] R. Zhang, Y.-C. Liang, C. C. Chai, and S. Cui, "Optimal beamforming for two-way multi-antenna relay channel with analogue network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 699–712, June 2009.

[11] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J*, vol. 54, no. 8, pp. 1355 –1387, Oct. 1975.

[12] C. E. Shannon, "Communication theory of secrecy systems," *Bell system technical journal*, vol. 28, no. 4, pp. 656–715, Oct. 1949.

[13] G. Vernam, "Cipher printing telegraph systems for secret wire and radio telegraphic communications," *Transactions of the American Institute of Electrical Engineers*, vol. XLV, pp. 295–301, Jan. 1926.

[14] U. Maurer, *Communications and cryptography: two sides of one tapestry.* Kluwer Academic Publishers, 1994, Chap. 27 The strong secret key rate of discrete random triples.

[15] I. Csiszár and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[16] A. Thangaraj, S. Dihidar, A. R. Calderbank, S. W. McLaughlin, and J.-M. Merolla, "Applications of LDPC codes to the wiretap channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 8, pp. 2933–2945, Aug. 2007.

[17] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas–Part II: The MIMOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.

[18] ——, "Secure transmission with multiple antennas I: The MISOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3088–3104, Jul. 2010.

[19] T. Liu, V. Prabhakaran, and S. Vishwanath, "The secrecy capacity of a class of parallel Gaussian compound wiretap channels," in *Proc. IEEE International Symposium on Information (ISIT'08)*, Toronto, ON, CA, July 2008, pp. 116–120.

[20] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai, "Compound wiretap channels," *EURASIP J. Wireless Commun. and Netw., special issue on Wireless Physical Layer Security*, vol. 2009, no. 5, pp. 1–12, Mar. 2009.

[21] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180–2189, June 2008.

[22] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4961–4972, Aug. 2011.

[23] T. Liu and S. Shamai, "A note on the secrecy capacity of the multiple-antenna wiretap channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2547–2553, Jun. 2009.

[24] S. A. A. Fakoorian and A. L. Swindlehurst, "Full rank solutions for the MIMO Gaussian wiretap channel with an average power constraint," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2620–2631, May 2013.

[25] ——, "Optimal power allocation for GSVD-based beamforming in the MIMO Gaussian wiretap channel," in *Proc. IEEE International Symposium on Information Theory (ISIT'12)*, Cambridge, MA, USA, Jul. 2012, pp. 2321–2325.

[26] R. Bustin, R. Liu, H. V. Poor, and S. Shamai, "An MMSE approach to the secrecy capacity of the MIMO Gaussian wiretap channel," *EURASIP J. Wirel. Commun. Netw.*, vol. 2009, pp. 3:1–3:8, Mar. 2009.

[27] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[28] L. Zhang, R. Zhang, Y.-C. Liang, Y. Xin, and S. Cui, "On the relationship between the multi-antenna secrecy communications and cognitive radio communications," *IEEE Trans. Commun.*, vol. 58, no. 6, pp. 1877–1886, June 2010.

[29] J. Huang and A. Swindlehurst, "Robust secure transmission in MISO channels based on worst-case optimization," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1696–1707, April 2012.

[30] A. Mukherjee and A. Swindlehurst, "Robust beamforming for security in MIMO wiretap channels with imperfect CSI," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 351–361, Jan. 2011.

[31] H.-M. Wang, M. Luo, X.-G. Xia, and Q. Yin, "Joint cooperative beamforming and jamming to secure AF relay systems with individual power constraint and no eavesdropper's CSI," *IEEE Signal Process. Lett.*, vol. 20, no. 1, pp. 39–42, Jan. 2013.

[32] S. Gerbracht, "Secrecy outage in MISO systems with partial channel information," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 2, pp. 704–716, April 2012.

[33] Q. Li and W.-K. Ma, "Spatially selective artificial-noise aided transmit optimization for MISO multi-Eves secrecy rate maximization," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2704–2717, May 2013.

[34] X. Zhou and M. McKay, "Secure transmission with artificial noise over fading channels: achievable rate and optimal power allocation," *IEEE Trans. Veh. Technol.*, vol. 59, no. 8, pp. 3831–3842, Oct. 2010.

[35] J. Li, A. P. Petropulu, and S. Weber, "On cooperative relaying schemes for wireless physical layer security," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4985–4997, Oct. 2011.

[36] G. Zheng, L.-C. Choo, and K.-K. Wong, "Optimal cooperative jamming to enhance physical layer security using relays," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1317–1322, 2011.

[37] J. Barros and M. Rodrigues, "Secrecy capacity of wireless channels," in *Proc. IEEE International Symposium on Information Theory (ISIT'06)*, Seattle, Washington, USA, July 2006, pp. 356–360.

[38] P. K. Gopala, L. Lai, and H. El-Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4687–4698, Oct. 2008.

[39] Y. Liang, H. Poor, and S. Shamai, "Secure communication over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2470–2492, June 2008.

[40] S. Ma, M. Hong, E. Song, X. Wang, and D. Sun, "Outage constrained robust secure transmission for MISO wiretap channels," *IEEE Trans. Wireless. Commun.*, vol. 13, no. 10, pp. 5558–5570, Oct. 2014.

[41] Z. Chu, H. Xing, M. Johnston, and S. Le Goff, "Secrecy rate optimizations for a MISO secrecy channel with multiple multi-antenna eavesdroppers," *to appear in IEEE Trans. Wireless Commun.*, 2015.

[42] O. Gungor, J. Tan, C. Koksal, H. El-Gamal, and N. Shroff, "Secrecy outage capacity of fading channels," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5379–5397, Sept. 2013.

[43] L. Lai and H. El Gamal, "The relay–eavesdropper channel: Cooperation for secrecy," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 4005–4019, Sept. 2008.

[44] J. Zhang and M. C. Gursoy, "Collaborative relay beamforming for secrecy," in *Proc. IEEE International Conference on Communications (ICC'10)*, Cape Town, ZA, May 2010, pp. 1–5.

[45] ——, "Relay beamforming strategies for physical-layer security," in *Proc. IEEE Annual Conference on Information Sciences and Systems (CISS'10)*, Princeton, NJ, USA, Mar. 2010, pp. 1–6.

[46] Y. Yang, Q. Li, W.-K. Ma, J. Ge, and P. Ching, "Cooperative secure beamforming for AF relay networks with multiple eavesdroppers," *IEEE Signal Process. Lett.*, vol. 20, no. 1, pp. 35–38, Jan. 2013.

[47] C. Jeong and I.-M. Kim, "Optimal power allocation for secure multicarrier relay systems," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5428–5442, Nov. 2011.

[48] X. Wang, K. Wang, and X.-D. Zhang, "Secure relay beamforming with imperfect channel side information," *IEEE Trans. Veh. Technol.*, vol. 62, no. 5, pp. 2140–2155, June 2013.

[49] Q. Li and W.-K. Ma, "Optimal and robust transmit designs for MISO channel secrecy by semidefinite programming," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3799–3812, Aug. 2011.

[50] S. Luo, J. Li, and A. Petropulu, "Uncoordinated cooperative jamming for secret communications," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 7, pp. 1081–1090, July 2013.

[51] Q. Li, Y. Yang, W.-K. Ma, M. Lin, J. Ge, and J. Lin, "Robust cooperative beamforming and artificial noise design for physical-layer secrecy in AF multi-antenna multi-relay networks," *IEEE Trans. Signal Process.*, vol. 63, no. 1, pp. 206–220, Jan. 2015.

[52] I. Krikidis, J. Thompson, and S. Mclaughlin, "Relay selection for secure cooperative networks with jamming," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 5003–5011, Oct. 2009.

[53] Z. Ding, K. K. Leung, D. L. Goeckel, and D. Towsley, "Opportunistic relaying for secrecy communications: Cooperative jamming vs. relay chatting," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1725–1729, Jun. 2011.

[54] S. Vishwakarma and A. Chockalingam, "Amplify-and-forward relay beamforming for secrecy with cooperative jamming and imperfect CSI," in *Proc. IEEE International Conference on Communications (ICC'13)*, Budapest, HU, June 2013, pp. 3047–3052.

[55] Y. Yang, Q. Li, W.-K. Ma, J. Ge, and M. Lin, "Optimal joint cooperative beamforming and artificial noise design for secrecy rate maximization in AF relay networks," in *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC'13)*, Darmstadt, DE, Jun. 2013, pp. 360–364.

[56] J. Huang and A. L. Swindlehurst, "Cooperative jamming for secure communications in MIMO relay networks," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4871–4884, Oct. 2011.

[57] L. Liu, R. Zhang, and K.-C. Chua, "Secrecy wireless information and power transfer with MISO beamforming," *IEEE Trans. Signal Process.*, vol. 62, no. 7, pp. 1850–1863, April 2014.

[58] D. W. K. Ng, E. S. Lo, and R. Schober, "Robust beamforming for secure communication in systems with wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4599–4615, Aug. 2014.

[59] M. Khandaker and K.-K. Wong, "Masked beamforming in the presence of energy-harvesting eavesdroppers," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 1, pp. 40–54, Jan. 2015.

[60] Q. Li, W.-K. Ma, and A.-C. So, "Robust artificial noise-aided transmit optimization for achieving secrecy and energy harvesting," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'14)*, Florence, IT, May 2014, pp. 1596–1600.

[61] B. Zhu, J. Ge, Y. Huang, Y. Yang, and M. Lin, "Rank-two beamformed secure multicasting for wireless information and power transfer," *IEEE Signal Process. Lett.*, vol. 21, no. 2, pp. 199–203, Feb. 2014.

[62] Q. Shi, W. Xu, J. Wu, E. Song, and Y. Wang, "Secure beamforming for MIMO broadcasting with wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2841–2853, May 2015.

[63] D. W. K. Ng, L. Xiang, and R. Schober, "Multi-objective beamforming for secure communication in systems with wireless information and power transfer," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*, London, UK, Sept. 2013, pp. 7–12.

[64] B. Fang, Z. Qian, W. Zhong, and W. Shao, "AN-aided secrecy precoding for SWIPT in cognitive MIMO broadcast channels," *IEEE Commun. Lett.*, vol. 19, no. 9, pp. 1632–1635, Sept. 2015.

[65] J. Zhang, C. Yuen, C.-K. Wen, S. Jin, K.-K. Wong, and H. Zhu, "Large system secrecy rate analysis for SWIPT MIMO wiretap channels," *to appear in IEEE Trans. Inf. Forensics Security*, 2015.

[66] Q. Li, Q. Zhang, and J. Qin, "Secure relay beamforming for simultaneous wireless information and power transfer in non-regenerative relay networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2462–2467, Jun. 2014.

[67] D. W. K. Ng, R. Schober, and H. Alnuweiri, "Secure layered transmission in multicast systems with wireless information and power transfer," in *Proc. IEEE International Conference on Communications (ICC'14)*, Sydney, NSW, AUS, Jun. 2014, pp. 5389–5395.

[68] Z. Chu, Z. Zhu, M. Johnston, and S. Le Goff, "Simultaneous wireless information power transfer for MISO secrecy channel," *to appear in IEEE Trans. Veh. Technol.*, 2015, available online at arXiv:1503.06437.

[69] R. Feng, Q. Li, Q. Zhang, and J. Qin, "Robust secure transmission in MISO simultaneous wireless information and power transfer system," *IEEE Trans. Veh. Technol.*, vol. 64, no. 1, pp. 400–405, Jan. 2015.

[70] Q. Zhang, X. Huang, Q. Li, and J. Qin, "Cooperative jamming aided robust secure transmission for wireless information and power transfer in MISO channels," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 906–915, Mar. 2015.

[71] H. Ju and R. Zhang, "Throughput maximization in wireless powered communication networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 418–428, Jan. 2014.

[72] W. Liu, X. Zhou, S. Durrani, and P. Popovski, "Secure communication with a wireless-powered friendly jammer," *to appear in IEEE Trans. Wireless Commun.*, 2015, available online at arXiv:1412.0349.

[73] M. Zhao, X. Wang, and S. Feng, "Joint power splitting and secure beamforming design in the multiple non-regenerative wireless-powered relay networks," *IEEE Commun. Lett.*, vol. 19, no. 9, pp. 1540–1543, Sept. 2015.

[74] M. Zhao, S. Feng, Y. Liu, X. Wang, M. Zhang, and H. Fu, "Joint power splitting and secure beamforming design in the wireless-powered untrusted relay networks," *to apprear in Proc. IEEE Global Communications Conference (GLOBECOM'15)*, 2015, available online at arXiv 1504.00770.

[75] X. Chen, J. Chen, and T. Liu, "Secure wireless information and power transfer in large-scale MIMO relaying systems with imperfect CSI," in *Proc. IEEE Global Communications Conference (CLOBECOM'14)*, Austin, TX, USA, Dec. 2014, pp. 4131–4136.

[76] H. Koorapaty, A. Hassan, and S. Chennakeshu, "Secure information transmission for mobile radio," in *Proc. IEEE International Symposium on Information (ISIT'98)*, Cambridge, MA,USA, Aug. 1998, p. 381.

[77] ——, "Secure information transmission for mobile radio," *IEEE Communications Letters*, vol. 4, no. 2, pp. 52–55, Feb. 2000.

[78] W. Liao, T. Chang, W. Ma, and C. Chi, "Qos-based transmit beamforming in the presence of eavesdroppers: An artificial-noise-aided approach," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1202–1216, Mar. 2011.

[79] K. Khalil, O. O. Koyluoglu, H. El-Gamal, and M. Youssef, "Opportunistic secrecy with a strict delay constraint," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4700–4709, Nov. 2013.

[80] A. Khisti, A. Tchamkerten, and G. W. Wornell, "Secure broadcasting over fading channels," *Trans. Inf. Theory*, vol. 54, no. 6, pp. 2453–2469, June 2008.

[81] L. Liu, R. Zhang, and K. Chua, "Wireless information transfer with opportunistic energy harvesting," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 288–300, Jan. 2013.

[82] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, May 2013.

[83] R. T. Rockafellar, *Convex Analysis.* Princeton Univ. Press, 1997.

[84] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge Univ. Press, 2004.

[85] R. Liu, T. Liu, H. Poor, and S. Shamai, "Multiple-input multiple-output gaussian broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4215–4227, Sep. 2010.

[86] X. Tang, R. Liu, P. Spasojevic, and H. V. Poor, "The Gaussian wiretap channel with a helping interferer," in *Proc. IEEE International Symposium on Information (ISIT'08)*, Toronto, ON, CA, July 2008, pp. 389–393.

[87] Z. Chu, K. Cumanan, Z. Ding, M. Johnston, and S. Le Goff, "Secrecy rate optimizations for a MIMO secrecy channel with a cooperative jammer," *IEEE Trans. Veh. Technol.*, vol. 64, no. 5, pp. 1833–1847, May 2015.

[88] K. Cumanan, Z. Ding, B. Sharif, G. Y. Tian, and K. K. Leung, "Secrecy rate optimizations for a MIMO secrecy channel with a multiple-antenna eavesdropper," *IEEE Trans. Veh. Technol.*, vol. 63, no. 4, pp. 1678–1690, May 2014.

[89] B. He, X. Zhou, and T. D. Abhayapala, "Wireless physical layer security with imperfect channel state information: A survey," *ZTE Commun.*, vol. 11, no. 3, pp. 11–19, Sept. 2013.

[90] A. Chorti, S. M. Perlaza, Z. Han, and H. V. Poor, "On the resilience of wireless multiuser networks to passive and active eavesdroppers," *IEEE J. Sel. Areas Commun*, vol. 31, no. 9, pp. 1850–1863, Sept. 2013.

[91] A. J. Laub, *Matrix Analysis for Scientists and Engineers.* SIAM, 2005.

[92] A. Charnes and W. W. Cooper, "Programming with linear fractional functionals," *Naval Res. Logist. Quart.*, vol. 9, no. 3-4, pp. 181–186, Sept.-Dec. 1962.

[93] H. Xing, K.-K. Wong, Z. Chu, and A. Nallanathan, "Robust MIMO harvest-and-jam helpers and relaying for secret communications," 2015, available online at arXiv:1502.07066v1.

[94] Y. Huang and D. P. Palomar, "Rank-constrained separable semidefinite programming with applications to optimal beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 664–678, Feb. 2010.

[95] Z.-Q. Luo, J. F. Sturm, and S. Zhang, "Multivariate nonnegative quadratic mappings," *SIAM J. Optim.*, vol. 14, no. 4, pp. 1140–1162, 2004.

[96] H. Wang, S. Ma, T.-S. Ng, and H. V. Poor, "A general analytical approach for opportunistic cooperative systems with spatially random relays," *IEEE Trans. Wireless Commun.*, vol. 10, no. 12, pp. 4122–4129, Dec. 2011.

[97] E. Karipidis, N. D. Sidiropoulos, and Z.-Q. Luo, "Far-field multicast beamforming for uniform linear antenna arrays," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 4916–4927, Oct. 2007.

[98] L. R. Varshney, "Transporting information and energy simultaneously," in *Proc. IEEE International Symposium on Information Theory (ISIT'08)*, Toronto, ON, CA, July 2008, pp. 1612–1616.

[99] L. Liu, R. Zhang, and K.-C. Chua, "Wireless information and power transfer: A dynamic power splitting approach," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3990–4001, Sept. 2013.

[100] S. Timotheou, I. Krikidis, G. Zheng, and B. Ottersten, "Beamforming for MISO interference channels with QoS and RF energy transfer," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2646–2658, May 2014.

[101] A. Nasir, X. Zhou, S. Durrani, and R. Kennedy, "Relaying protocols for wireless energy harvesting and information processing," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3622–3636, July 2013.

[102] K. Ishibashi, "Dynamic harvest-and-forward: New cooperative diversity with RF energy harvesting," in *Proc. IEEE International Conference on Wireless Communications and Signal Processing (WCSP'14)*, Hefei, CHN, Oct. 2014, pp. 1–5.

[103] Z. Ding, S. Perlaza, I. Esnaola, and H. Poor, "Power allocation strategies in energy harvesting wireless cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 846–860, Feb. 2014.

[104] Q. Shi, L. Liu, W. Xu, and R. Zhang, "Joint transmit beamforming and receive power splitting for MISO SWIPT systems," *IEEE Trans. Wireless. Commun.*, vol. 13, no. 6, pp. 3269–3280, June 2014.

[105] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[106] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[107] J. Zhu, R. Schober, and V. K. Bhargava, "Secure transmission in multicell massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4766–4781, Sept. 2014.

[108] H. Q. Ngo, H. A. Suraweera, M. Matthaiou, and E. G. Larsson, "Multipair full-duplex relaying with massive arrays and linear processing," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1721–1737, Sept. 2014.

[109] X. Chen, L. Lei, H. Zhang, and C. Yuen, "Large-scale MIMO relaying techniques for physical layer security: AF or DF?" *IEEE Trans. Wireless Commun.*, vol. 14, no. 9, pp. 5135–5146, Sept. 2015.

[110] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Jun. 2015.

[111] A. Jeffrey and D. Zwillinger, *Table of Integrals, Series, and Products*. Academic Press, 2007.

[112] W. Ai and S. Zhang, "Strong duality for the CDT subproblem: A necessary and sufficient condition," *SIAM J. Optim.*, vol. 19, no. 4, pp. 1735–1756, Feb. 2009.

[113] Y. Eldar, A. Ben-Tal, and A. Nemirovski, "Robust mean-squared error estimation in the presence of model uncertainties," *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 168–181, Jan. 2005.

[114] O. Klesov, *Limit Theorems for Multi-Indexed Sums of Random Variables*, ser. Probability Theory and Stochastic Modelling. Berlin, Germany: Springer, 2014, vol. 71.

List of Publications

Journal Publications

H. Xing, L. Liu, and R. Zhang, "Secrecy wireless information and power transfer in fading wiretap channel," to appear in *IEEE Transactions on Vehicular Technology*, 2015, available online at arXiv:1408.1987.

Z. Chu, **H. Xing**, M. Johnston, and S. Le Goff, "Secrecy rate optimizations for a MISO secrecy channel with multiple multi-antenna eavesdroppers," to appear in *IEEE Transactions on Wireless Communications*, 2015.

H. Xing, K-K. Wong, C. Zheng, and A. Nallanathan, "To harvest and jam: a paradigm of self-sustaining friendly jammers for secure AF relaying," *IEEE Transactions on Signal Processing*, vol. 63, no.24, pp. 6616-6631, Dec. 2015.

H. Zhang, **H. Xing**, J. Cheng, A. Nallanathan, and V.C.M. Leung, "Secure resource allocation for OFDMA two-way relay wireless sensor networks without and with cooperative jamming," to appear in *IEEE Transactions on Industrial Informatics*, 2015.

H. Xing, K-K. Wong, A. Nallanathan, and R. Zhang, "Wireless powered cooperative jamming for secrecy multi-AF relaying networks," submitted to *IEEE Transactions on Wireless Communications*, available online at arXiv:1511.03705.

C. Cumanan, **H. Xing**, P. Xu, G. Zheng, X. Dai, A. Nallanathan, Z. Ding, and G.K. Karagiannidis, "Physical layer security jamming: theoretical limits and practical designs in wireless networks," submitted to *IEEE Wireless Communications*

Magazine.

Conference Publications

H. Zhang, **H. Xing**, X. Chu, A. Nallanathan, W. Zheng, and X. Wen, "Secure resource allocation for OFDM two-way relay networks," in *Proc. IEEE Global Communications Conference (GlOBECOM)*, Anaheim, CA, USA, Dec. 2012, pp. 3649-3654.

H. Xing, H. Zhang, Z. Ding, X. Chu, and A. Nallanathan, "Secure resource allocation for OFDM two-way relay networks via cooperative jamming," accepted for publishing in *Proc. IEEE International Conference on Communications and Networking in China (CHINACOM)*, Kunming, China, Aug. 2012. **(Invited)**

H. Xing, L. Liu, and R. Zhang, "Secrecy wireless information and power transfer in fading wiretap channel," in *IEEE International Conference on Communications (ICC)*, Sydney, AUS, Jun. 2014, pp. 5402-5407.

H. Xing, C. Zheng, Z. Ding, and A. Nallanathan, "Harvest-and-Jam: improving security for wireless energy harvesting cooperative networks," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Austin, TX, USA, Dec. 2014, pp. 3145-3150. **(Best 50 Paper of GLOBECOM14)**

O. Holland, S. Ping, N. Sastry, **H. Xing** et al., "Some initial results and observations from a series of trials within the Ofcom TV white spaces pilot," in *Proc. IEEE Vehicular Technology Conference (VTC Spring)*, Glasgow, UK, May 2015, pp. 1-7.

H. Xing, K-K. Wong, and A. Nallanathan, "Secure wireless energy harvesting-enabled AF-relaying SWIPT networks," in *IEEE International Conference on Communications (ICC)*, London, UK, Jun. 2015, pp. 2307-2312.